

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2003, 13:22:16 ; Search time 14650 Seconds

(without alignments)
17411.025 Million cell updates/sec

Title: US-09-845-020a-5

Perfect score: 6235
Sequence: 1 gatcaacttgagacagtagt.....tgcacatcccccacagacct 6235Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl1.*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.yl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.scs:*
28: em.un:*
29: em.yl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6235	100.0	6235	6	ARI56464	ARI56464 Sequence
2	6235	100.0	6235	6	AX250679	AX250679 Sequence
3	6235	100.0	6679	6	ARI56461	ARI56461 Sequence
4	6235	100.0	6679	6	AX250678	AX250678 Sequence
5	6051.2	97.1	227857	6	AC090844	AC090844 Homo sapi
6	6032	96.7	122132	9	AC102729	AC102729 Homo sapi
7	5671.6	91.0	100490	2	AC007776	AC007776 Homo sapi
8	4368.4	70.1	221484	2	AC019095	AC019095 Homo sapi
9	2834	45.5	2834	6	ARI56465	ARI56465 Sequence
10	1612.4	25.9	5527	9	AF388025	AF388025 Homo sapi
11	1424.8	22.9	221484	2	AC019095	AC019095 Homo sapi
12	1252	20.1	1252	6	ARI56466	ARI56466 Sequence
13	1189.4	19.1	110779	9	BX470173	BX470173 Human DNA
14	1187.8	19.1	51015	9	AL732292	AL732292 Human DNA
15	1129.8	18.1	196806	9	AC104212	AC104212 Homo sapi
16	1109	17.8	189036	2	AL645608	AL645608 Homo sapi
17	1108.8	17.8	59925	9	AC092543	AC092543 Homo sapi
18	1084.4	17.4	93048	9	AC108148	AC108148 Homo sapi
19	1081.8	17.4	156387	2	BX537114	BX537114 Homo sapi
20	1081.2	17.3	155334	2	AC015822	AC015822 Homo sapi
21	1081.2	17.3	162036	2	AC015773	AC015773 Homo sapi
22	1080.2	17.3	144444	9	HSJ365119	AL078463 Human DNA
23	1071.2	17.2	142498	2	AC023980	AC023980 Homo sapi
24	1071.2	17.2	145787	9	AC099499	AC099499 Homo sapi
25	1051.8	16.9	184942	9	AC097262	AC097262 Homo sapi
26	1046.6	16.8	166720	9	AC112180	AC112180 Homo sapi
27	1046.6	16.8	192518	2	AC027745	AC027745 Homo sapi
28	1046.6	16.8	203831	9	AC093295	AC093295 Homo sapi
29	1045.4	16.8	110000	2	AL831785-7	Continuation (8 of
30	1035.4	16.6	107908	9	AL360155	AL360155 Human DNA
31	1035.4	16.6	157800	2	AL357074	AL357074 Homo sapi
32	1035.4	16.6	159790	2	AC013613	AC013613 Homo sapi
33	1029	16.5	169063	9	AC138126	AC138126 Homo sapi
34	1029	16.5	177299	9	AC073544	AC073544 Homo sapi
35	1028.4	16.5	202488	2	AC032041	AC032041 Homo sapi
36	1025.2	16.4	150972	9	AL512380	AL512380 Human DNA
37	1024	16.4	227430	2	AC091590	AC091590 Homo sapi
38	1017.6	16.3	42969	9	AC004197	AC004197 Homo sapi
39	1017.6	16.3	43901	2	AC004198	AC004198 Homo sapi
40	1017.6	16.3	117327	9	AL663093	AL663093 Human DNA
41	1016.6	16.3	141980	9	AL450425	AL450425 Human DNA
42	1016.6	16.3	162299	2	AC016429	AC016429 Homo sapi
43	1013.8	16.3	177883	2	AC138471	AC138471 Homo sapi
44	1008.6	16.2	145456	9	AL513366	AL513366 Human DNA
45	1007.2	16.2	62988	9	AL731680	AL731680 Human DNA

ALIGNMENTS

RESULT 1
LOCUS ARI56464
DEFINITION Sequence 5 from patent US 6242218.
ACCESSION ARI56464
VERSION ARI56464.1 GI:15125168
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6235)
AUTHORS Treco,D.A., Heartlein,M.W. and Seiden,R.F.
TITLE Genomic sequences for protein production and delivery
JOURNAL Patent: US 6242218-A 5 05-JUN-2001;
FEATURES Location/Qualifiers

source 1..6235
/organism="unknown"
BASE COUNT 1586 a 1610 c 1708 g 1331 t
ORIGIN

Query Match 100.0%; Score 6235; DB 6; Length 6235;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGCTTGAGACAGTACGTTCAAGACACCGCTGGGACGATAGGGAGACTGTCTACG 60
DB 1 GATCAGCTTGAGACAGTACGTTCAAGACACCGCTGGGACGATAGGGAGACTGTCTACG 60
QY 61 AAAATCAAAAAATTATGGCGGGGCGATGGCTGACCTGTATCCCTGAACCTTTGGG 120
DB 61 AAAATCAAAAAATTATGGCGGGGCGATGGCTGACCTGTATCCCTGAACCTTTGGG 120
QY 121 ACATCAAGGCAAGTGGATCACTTGAAGTCAAGAGTTCAGACTAGCTGGCCCAACATGGT 180
DB 121 ACATCAAGGCAAGTGGATCACTTGAAGTCAAGAGTTCAGACTAGCTGGCCCAACATGGT 180
QY 181 GAAACCCATATCTCCACTAAAAAATACAAAAATTAGCCAGCATGGTGGCAGGACCTGTGA 240
DB 181 GAAACCCATATCTCCACTAAAAAATACAAAAATTAGCCAGCATGGTGGCAGGACCTGTGA 240
QY 241 ATCCCGGCTACTCAGAGGCGTGAAGGAGGAGATCACTTGAACCCAGGAGGCGGAGTTG 300
DB 241 ATCCCGGCTACTCAGAGGCGTGAAGGAGGAGATCACTTGAACCCAGGAGGCGGAGTTG 300
QY 301 CAGTGAAGTGAATCAGACACCACTGCATCCAGCTGGGTGACAGAGCAAGACTGTATCTC 360
DB 301 CAGTGAAGTGAATCAGACACCACTGCATCCAGCTGGGTGACAGAGCAAGACTGTATCTC 360
QY 361 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
DB 361 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
QY 421 CTACTCAGAGGCTGAGGTGGAGGATCACTTGAACCTGGGGCAGTCAAGGCTACAGTGA 480
DB 421 CTACTCAGAGGCTGAGGTGGAGGATCACTTGAACCTGGGGCAGTCAAGGCTACAGTGA 480
QY 481 GCCAAGATCATGCCACTACACTCCAGCCTGGGCAACAGAGAGAGACCCTGTCTAAAAA 540
DB 481 GCCAAGATCATGCCACTACACTCCAGCCTGGGCAACAGAGAGAGACCCTGTCTAAAAA 540
QY 541 AATATATATATATAAAGAAAAAACAAGCTGTATATGCTCTGCTGCATACATACACT 600
DB 541 AATATATATATATAAAGAAAAAACAAGCTGTATATGCTCTGCTGCATACATACACT 600
QY 601 ATGATATAGTTGGCAAACTCAAGATCAGATAGTCAATTTTAAAGCTTGTGGGCGT 660
DB 601 ATGATATAGTTGGCAAACTCAAGATCAGATAGTCAATTTTAAAGCTTGTGGGCGT 660
QY 661 ATGGTCTCTGCACATCACTCTGCTGTCTTCTAGCACAAAAAGCAGCTATTAACAT 720
DB 661 ATGGTCTCTGCACATCACTCTGCTGTCTTCTAGCACAAAAAGCAGCTATTAACAT 720
QY 721 ACATATATGAATTTTATATGACATGAGATTTGAATTCATATGATTTTACATTTTAT 780
DB 721 ACATATATGAATTTTATATGACATGAGATTTGAATTCATATGATTTTACATTTTAT 780
QY 781 AAAATATATCTTTTAAAAAATTTTCCCTAACATTTTAAAGTGAAGGCGGCGCAGCGC 840
DB 781 AAAATATATCTTTTAAAAAATTTTCCCTAACATTTTAAAGTGAAGGCGGCGCAGCGC 840
QY 841 GCCATGCTCAGCGCTGTAAATTCAGACATTTGGAGGCTGAGGTGGCAGATCACTTGAG 900
DB 841 GCCATGCTCAGCGCTGTAAATTCAGACATTTGGAGGCTGAGGTGGCAGATCACTTGAG 900
QY 901 ATCAACAGTTGAGAGACCAAGCTGGCCAAATAGCAAAACCATTTCTACTAAAAATAAA 960
DB 901 ATCAACAGTTGAGAGACCAAGCTGGCCAAATAGCAAAACCATTTCTACTAAAAATAAA 960

QY 961 AAAATTAGCTGGGCACTAGTGGTGCACACCTGTGATCCCACTACTTGGAGGCTGAGGCA 1020
DB 961 AAAATTAGCTGGGCACTAGTGGTGCACACCTGTGATCCCACTACTTGGAGGCTGAGGCA 1020
QY 1021 GGAGAAATGCTTGAACCTGGGAAGCGGAGGTGGAGTGGAGCCAACTATGCTGACATGGAC 1080
DB 1021 GGAGAAATGCTTGAACCTGGGAAGCGGAGGTGGAGTGGAGCCAACTATGCTGACATGGAC 1080
QY 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAATAAGTGTAAAAAGCAT 1140
DB 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAATAAGTGTAAAAAGCAT 1140
QY 1141 TCCATATTCAGTGTACATCACTGTATCAATCTAGCTGCTGCTGCTGAGGCAAT 1200
DB 1141 TCCATATTCAGTGTACATCACTGTATCAATCTAGCTGCTGCTGCTGAGGCAAT 1200
QY 1201 ACCGTGAGAAATGAGATTTGCTGTGACAGGAGCATACATTTCCATTAATCACTAGACAT 1260
DB 1201 ACCGTGAGAAATGAGATTTGCTGTGACAGGAGCATACATTTCCATTAATCACTAGACAT 1260
QY 1261 ACCAAGTTGCCATCCAAAGAGGTTTTTTTTTACAACTAGACTCCCCAGCAACAAT 1320
DB 1261 ACCAAGTTGCCATCCAAAGAGGTTTTTTTTTACAACTAGACTCCCCAGCAACAAT 1320
QY 1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTCTAAGCCCACTACAGATGAAAAACAGA 1380
DB 1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTCTAAGCCCACTACAGATGAAAAACAGA 1380
QY 1381 AGTGGAGGAGGAGCTGCACGCCCCCTTAACCAATGAAGAAATCCCTGTAGAGCTTCT 1440
DB 1381 AGTGGAGGAGGAGCTGCACGCCCCCTTAACCAATGAAGAAATCCCTGTAGAGCTTCT 1440
QY 1441 GGATGCTGGAAGATGAATTAACGGGGGCTCTGAGAGCTGCCCTGTACATCACTGTG 1500
DB 1441 GGATGCTGGAAGATGAATTAACGGGGGCTCTGAGAGCTGCCCTGTACATCACTGTG 1500
QY 1501 ACTTCTGAGCCTCCAGTCCATCTCAGGCCCATGTGTATGGCCAGTATATAGGCCCT 1560
DB 1501 ACTTCTGAGCCTCCAGTCCATCTCAGGCCCATGTGTATGGCCAGTATATAGGCCCT 1560
QY 1561 CACTCTGTTTGGCTTTATATCTCCCATGTGGGGCTGAAGTGTGATTAAGCGCTTAT 1620
DB 1561 CACTCTGTTTGGCTTTATATCTCCCATGTGGGGCTGAAGTGTGATTAAGCGCTTAT 1620
QY 1621 TCAAGATGTACAGCTTCTTACAGAAAGTATGTCTACAGAAACAGCAGGGCTTGCA 1680
DB 1621 TCAAGATGTACAGCTTCTTACAGAAAGTATGTCTACAGAAACAGCAGGGCTTGCA 1680
QY 1681 AGATGATCTAATCTGAATCTTACCTGGCTCAGCCACAGCTAGTCTGTGATCTTGAAAC 1740
DB 1681 AGATGATCTAATCTGAATCTTACCTGGCTCAGCCACAGCTAGTCTGTGATCTTGAAAC 1740
QY 1741 AAGTTTTTCACTTCTGAGAGCCATCCCTTGGCTACAAACACAGATTTGTTAGACAGA 1800
DB 1741 AAGTTTTTCACTTCTGAGAGCCATCCCTTGGCTACAAACACAGATTTGTTAGACAGA 1800
QY 1801 TGAATATGACAAATGCTTACACCTGTATCTCAGACACTTTGGAGGCGCAAGGCGGGTGG 1860
DB 1801 TGAATATGACAAATGCTTACACCTGTATCTCAGACACTTTGGAGGCGCAAGGCGGGTGG 1860
QY 1861 ATGGCTGAGGCTGAGAGGTGACAGCATGCCGGGAGTCTTCACAGCCCTGTGCTGCTC 1920
DB 1861 ATGGCTGAGGCTGAGAGGTGACAGCATGCCGGGAGTCTTCACAGCCCTGTGCTGCTC 1920
QY 1921 GGGCCTCTCTCTGCTGGGCTCCCACTTGGTGGCACTTGAAGAGGAGGCTTACAGCCACG 1980
DB 1921 GGGCCTCTCTCTGCTGGGCTCCCACTTGGTGGCACTTGAAGAGGAGGCTTACAGCCACG 1980
QY 1981 CTGCACGTGTGGAGCCCTTTCTGGGCTGGCCAAAGGCGGCTCCCTCAGCTTGC 2040
DB 1981 CTGCACGTGTGGAGCCCTTTCTGGGCTGGCCAAAGGCGGCTCCCTCAGCTTGC 2040
QY 2041 AGGAGGTTGTGAGAGGAGGCTCAACAGCAACCGGGGCTGCCACAGGCGGCTTGGCGGC 2100

|||||
Db 2041 AGGAGGCTGTGAGAGGAGGCTCAAGAGAAACGGGGCTGCGACGGCGCTGGGGG 2100
Qy 2101 CAGGTGAGTTCGGGGTGGGGCTGGCGGGGCCCCGACATCGAGAGCGGGCCAG 2160
Db 2101 CAGGTGAGTTCGGGGTGGGGCTGGCGGGGCCCCGACATCGAGAGCGGGCCAG 2160
Qy 2161 CCCTGCGAGGCCCCGGGCAATGAGAGGCTTAGACACCGGGGCGAGCGGCTGGAGAGGT 2220
Db 2161 CCCTGCGAGGCCCCGGGCAATGAGAGGCTTAGACACCGGGGCGAGCGGCTGGAGAGGT 2220
Qy 2221 ACTGGGTGCCCCCAGACAGTGCAGCGCCGCGCGCTGTGCTGTGATTTCTCACTGGGC 2280
Db 2221 ACTGGGTGCCCCCAGACAGTGCAGCGCCGCGCGCTGTGCTGTGATTTCTCACTGGGC 2280
Qy 2281 CTTAGCAGCCTTCCGGGGGCGAGGGCTGGGACCTGAGCGCCGATGCTGAGCCCTCC 2340
Db 2281 CTTAGCAGCCTTCCGGGGGCGAGGGCTGGGACCTGAGCGCCGATGCTGAGCCCTCC 2340
Qy 2341 CTTGATGAGGCTCTGTGCGGGCCGAGGCTCCCGAGAGCAGACCCCTGTCTCAGAG 2400
Db 2341 CTTGATGAGGCTCTGTGCGGGCCGAGGCTCCCGAGAGCAGACCCCTGTCTCAGAG 2400
Qy 2401 CGCCGAGTCCCATCGACCGAGGCTGAGAAAGTGGGCGGCGACGGGACCTGG 2460
Db 2401 CGCCGAGTCCCATCGACCGAGGCTGAGAAAGTGGGCGGCGACGGGACCTGG 2460
Qy 2461 CAGGACCTACCCCTGCGAGCCCTGTGTGGAATCCTACTGGGTGAAGCCAGCTGGCTCT 2520
Db 2461 CAGGACCTACCCCTGCGAGCCCTGTGTGGAATCCTACTGGGTGAAGCCAGCTGGCTCT 2520
Qy 2521 GAGCTGATGAGAGCTTGAGAACCTTATGTCTAGCTCAGAGGATCGTAAATACCAAT 2580
Db 2521 GAGCTGATGAGAGCTTGAGAACCTTATGTCTAGCTCAGAGGATCGTAAATACCAAT 2580
Qy 2581 CAGCACCCTGTGTCTAGCTCAGGCTGTGTGGAATCAGCAATCAGCACTGTATCTAGCT 2640
Db 2581 CAGCACCCTGTGTCTAGCTCAGGCTGTGTGGAATCAGCAATCAGCACTGTATCTAGCT 2640
Qy 2641 ACTGTGATGGGGCTTGAGAACCTTATGTCTAGCTCAGAGGATGTTAAATACCAATC 2700
Db 2641 ACTGTGATGGGGCTTGAGAACCTTATGTCTAGCTCAGAGGATGTTAAATACCAATC 2700
Qy 2701 GGCACCTGTATCTAGCTCAGAGTGTGTAACACCAATCAGCACTGTGTCTAGCTC 2760
Db 2701 GGCACCTGTATCTAGCTCAGAGTGTGTAACACCAATCAGCACTGTGTCTAGCTC 2760
Qy 2761 AGGCTATGTGAATGACCAATCGACAGTCTGTATCTGTGCTACTTTCAATGGGCATCCGT 2820
Db 2761 AGGCTATGTGAATGACCAATCGACAGTCTGTATCTGTGCTACTTTCAATGGGCATCCGT 2820
Qy 2821 GAAAGAGCACCAAGAGGCTTTGTGTGAGCAATTAAGCTTCTATCCTCTGGGTGACAGT 2880
Db 2821 GAAAGAGCACCAAGAGGCTTTGTGTGAGCAATTAAGCTTCTATCCTCTGGGTGACAGT 2880
Qy 2881 GGGGTGAGTCCGAAAGAGAGTCAAGCGAAGGAGGTAAGGGTGGGGCTTTTATAGAT 2940
Db 2881 GGGGTGAGTCCGAAAGAGAGTCAAGCGAAGGAGGTAAGGGTGGGGCTTTTATAGAT 2940
Qy 2941 TTGGGTAGGTAAAGAAATTAACAGTCAAAAGGGGTTTGTCTCTGGGGGCGAGAGTGG 3000
Db 2941 TTGGGTAGGTAAAGAAATTAACAGTCAAAAGGGGTTTGTCTCTGGGGGCGAGAGTGG 3000
Qy 3001 GGGGTGCGAAGGCTCTAGTGGGGTGTCTTTTGAAGCGAGATGAGCGAGGAAAAAGAGT 3060
Db 3001 GGGGTGCGAAGGCTCTAGTGGGGTGTCTTTTGAAGCGAGATGAGCGAGGAAAAAGAGT 3060
Qy 3061 TTTCAGAGTAAATCTCATTAATTAAGGCAAGGACCGGCATTTACACTCTTTTGTGTG 3120
Db 3061 TTTCAGAGTAAATCTCATTAATTAAGGCAAGGACCGGCATTTACACTCTTTTGTGTG 3120
Qy 3121 GAATGTATCATGTTAAAGTGGGCGAGGCAATTCATCTTTTGTGTGATCTTCAAGTAC 3180
|||||

|||||
Db 3121 GAATGTATCATGTTAAAGTGGGCGAGGCAATTCATCTTTTGTGTGATCTTCAAGTAC 3180
Qy 3181 TTTCAGGCAATCGGGCGTAAATATGTCAGTAAAGGGGATGCGATGCTTGGCTGGGCT 3240
Db 3181 TTTCAGGCAATCGGGCGTAAATATGTCAGTAAAGGGGATGCGATGCTTGGCTGGGCT 3240
Qy 3241 CAGAGGCTTGACAGCTACTGTGTGTTGGGCGCTTGGAGAAATGTTGTGTGACACTGTAT 3300
Db 3241 CAGAGGCTTGACAGCTACTGTGTGTTGGGCGCTTGGAGAAATGTTGTGTGACACTGTAT 3300
Qy 3301 CTAGTTAATCTAGTGGGAGCTGGAACCTTTGTGTCTAGCTCAGGATGTTAAACCA 3360
Db 3301 CTAGTTAATCTAGTGGGAGCTGGAACCTTTGTGTCTAGCTCAGGATGTTAAACCA 3360
Qy 3361 CCAATCAGCGGCGCTTCAAAACAGACCACTGGGCTCTACCAATGAGAGATGTTGGTGG 3420
Db 3361 CCAATCAGCGGCGCTTCAAAACAGACCACTGGGCTCTTACCAATGAGAGATGTTGGTGG 3420
Qy 3421 GGCAGATTAAGAAATAAAGCAGGCTCCGAGCGCAGAGTGGCAAGCGCACAGTCC 3480
Db 3421 GGCAGATTAAGAAATAAAGCAGGCTCCGAGCGCAGAGTGGCAAGCGCACAGTCC 3480
Qy 3481 CTATCCAAATATGAGCAGCTTTGTTTGTGCTGTTGCGATTAATCTTGTCTACTGTG 3540
Db 3481 CTATCCAAATATGAGCAGCTTTGTTTGTGCTGTTGCGATTAATCTTGTCTACTGTG 3540
Qy 3541 CTTTTGGGTCCACACTCTTTTATGACCTGTAAACATCCACGAAGTCTGCAAGCTTC 3600
Db 3541 CTTTTGGGTCCACACTCTTTTATGACCTGTAAACATCCACGAAGTCTGCAAGCTTC 3600
Qy 3601 ACTCCTGAAGCCTAAAGACCAAGAGCCAGCGGAGGAAATGAAACATCCGCGGCGCT 3660
Db 3601 ACTCCTGAAGCCTAAAGACCAAGAGCCAGCGGAGGAAATGAAACATCCGCGGCGCT 3660
Qy 3661 GCCTTAAGAGCTATTAACACTACCGGAGAGTCTGACCTTCACTCTGACGACGAGA 3720
Db 3661 GCCTTAAGAGCTATTAACACTACCGGAGAGTCTGACCTTCACTCTGACGACGAGA 3720
Qy 3721 CCAGAGCCCAACCAAGAAAGAACTGGAACATGGAACATCAGAGGAACAACT 3780
Db 3721 CCAGAGCCCAACCAAGAAAGAACTGGAACATGGAACATCAGAGGAACAACT 3780
Qy 3781 CCAGATGACCACTTAAAGAGCTTAACACTGACGAGGCTCCGCGCTCTCTTCTG 3840
Db 3781 CCAGATGACCACTTAAAGAGCTTAACACTGACGAGGCTCCGCGCTCTCTTCTG 3840
Qy 3841 AAGTCAGTGAACCAAGCACTCACAGTGTGGAACACAGCCAGAGTGTGAGATCAG 3900
Db 3841 AAGTCAGTGAACCAAGCACTCACAGTGTGGAACACAGCCAGAGTGTGAGATCAG 3900
Qy 3901 CTGGGCAACATGATTAATGCCCCTCTGTGCAAAAAAAATTAACAAAAATTTGGGG 3960
Db 3901 CTGGGCAACATGATTAATGCCCCTCTGTGCAAAAAAAATTAACAAAAATTTGGGG 3960
Qy 3961 AGCATGGTGGTCCGCTGGTGGTCCAGCTAACGGGAGGTAAGTGGGAGATGCT 4020
Db 3961 AGCATGGTGGTCCGCTGGTGGTCCAGCTAACGGGAGGTAAGTGGGAGATGCT 4020
Qy 4021 TGAGCCTGGAGAGTAAAGTCACTGAGTGTGTTTATACACAGCCCTCTAGGCTGG 4080
Db 4021 TGAGCCTGGAGAGTAAAGTCACTGAGTGTGTTTATACACAGCCCTCTAGGCTGG 4080
Qy 4081 GGCAGACTGAGACCCCTTTCCCTCCGCAAAAAAATTAACAAAAATGTAATAGAGT 4140
Db 4081 GGCAGACTGAGACCCCTTTCCCTCCGCAAAAAAATTAACAAAAATGTAATAGAGT 4140
Qy 4141 GCTGTATATGCTAGGCGCAGTGTGCTATGCTGTAAATCCAGCACTTTGGAAACCCAG 4200
Db 4141 GCTGTATATGCTAGGCGCAGTGTGCTATGCTGTAAATCCAGCACTTTGGAAACCCAG 4200
Qy 4201 GCGGGCGGCTCAAGCTTAAGGTCAGAGTGTGAGACAGGCTGCGCAACATGGAAGAGCC 4260
Db 4201 GCGGGCGGCTCAAGCTTAAGGTCAGAGTGTGAGACAGGCTGCGCAACATGGAAGAGCC 4260

QY	4261	ATCTCTTCTAAAAATACAAAATTAGCCGGCTGTGTGGGGGCGACTGGTGGAGCATGGCTGTAA	4320
Db	4261	ATCTCTTCTAAAAATACAAAATTAGCCGGCTGTGTGGGGGCGACTGGTGGAGCATGGCTGTAA	4320
QY	4321	TCCCGCATCTCAGGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGAGCGCGGCTTGC	4380
Db	4321	TCCCGCATCTCAGGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGAGCGCGGCTTGC	4380
QY	4381	AGTGAAGCCGAGATCGTGGCATTTGCACTCCACCACTCCAGCTGGGGCAACGAAGGCCAAA	4440
Db	4381	AGTGAAGCCGAGATCGTGGCATTTGCACTCCACCACTCCAGCTGGGGCAACGAAGGCCAAA	4440
QY	4441	CTCTGTCTTAAAAAAGTGCCTGCATATAAGAGGTGTGCATATCAATAG	4500
Db	4441	CTCTGTCTTAAAAAAGTGCCTGCATATAAGAGGTGTGCATATCAATAG	4500
QY	4501	TTGCCAGCGCAACATGTTTAAGATGTGGAGCTCCTGCCCTCATGTGTCCTTTAAAAACC	4560
Db	4501	TTGCCAGCGCAACATGTTTAAGATGTGGAGCTCCTGCCCTCATGTGTCCTTTAAAAACC	4560
QY	4561	CACCCCTCAAGGCGAGGTGCAATGGCTCTAAATCCAGACACTTTGGGAGCCGAG	4620
Db	4561	CACCCCTCAAGGCGAGGTGCAATGGCTCTAAATCCAGACACTTTGGGAGCCGAG	4620
QY	4621	GCGGGTGGATCACTGAGGTGAGAGATTGAGAGACAGGCTGACCAACAACTGGTGAAT	4680
Db	4621	GCGGGTGGATCACTGAGGTGAGAGATTGAGAGACAGGCTGACCAACAACTGGTGAAT	4680
QY	4681	CCGACCTTACTAAAAATACAAATTTAGATGACATGGTGTGCATGCTGTAAATCCAC	4740
Db	4681	CCGACCTTACTAAAAATACAAATTTAGATGACATGGTGTGCATGCTGTAAATCCAC	4740
QY	4741	CTACTTGGGAGGCTGAGGCGAGAAAATACATAGAAACGAGGAGCGGAGGTTGTAGTGA	4800
Db	4741	CTACTTGGGAGGCTGAGGCGAGAAAATACATAGAAACGAGGAGCGGAGGTTGTAGTGA	4800
QY	4801	CCGAGATGTCATTCATTCAGCTTCAGCCTGACCAATGAGCGAAACTCCATCTCAAAAAAC	4860
Db	4801	CCGAGATGTCATTCATTCAGCTTCAGCCTGACCAATGAGCGAAACTCCATCTCAAAAAAC	4860
QY	4861	AACAACAAAAAACCACCTCTCTACTCCAGGAGGCTGGGTACAGAGCTGGGCCACATCAGT	4920
Db	4861	AACAACAAAAAACCACCTCTCTACTCCAGGAGGCTGGGTACAGAGCTGGGCCACATCAGT	4920
QY	4921	GCAAGTCTCTAGCCACAGAGACTTAAGGCGGAGCTGCGAGAGCCGGGACAGATATAACGTG	4980
Db	4921	GCAAGTCTCTAGCCACAGAGACTTAAGGCGGAGCTGCGAGAGCCGGGACAGATATAACGTG	4980
QY	4981	TGTGAGATCAATGTGTGTGAGATCAGACGTCCTTGGCATTTGGTGAACACAGGAGGGCCCCA	5040
Db	4981	TGTGAGATCAATGTGTGTGAGATCAGACGTCCTTGGCATTTGGTGAACACAGGAGGGCCCCA	5040
QY	5041	AGCACAGAGATGGCCCCCATCCAGTCACACATCCACTTTCATTCAGAGATGTCTGT	5100
Db	5041	AGCACAGAGATGGCCCCCATCCAGTCACACATCCACTTTCATTCAGAGATGTCTGT	5100
QY	5101	CTTGGCAGCGTGGGGTAAATTAGACAGAGAGTGAACACTTCTGGGTGGTCACTCAGAC	5160
Db	5101	CTTGGCAGCGTGGGGTAAATTAGACAGAGAGTGAACACTTCTGGGTGGTCACTCAGAC	5160
QY	5161	TGCCCGAAGGCGAGGCTTGTGGCTGTGTGAATAACCTTCAGGCTTAAGGCCGGGACAGGTGGC	5220
Db	5161	TGCCCGAAGGCGAGGCTTGTGGCTGTGTGAATAACCTTCAGGCTTAAGGCCGGGACAGGTGGC	5220
QY	5221	TCACGCGCTGTATCCACAGACTTTTGGAGGCGGAGGCGGAGGTGATCAAGAGTTCAGAGA	5280
Db	5221	TCACGCGCTGTATCCACAGACTTTTGGAGGCGGAGGCGGAGGTGATCAAGAGTTCAGAGA	5280
QY	5281	TCGTGACATCTCTGGCTTAACAGCGGTGAACCCCGCTCTCTATAAAATACAAAAATTTGG	5340
Db	5281	TCGTGACATCTCTGGCTTAACAGCGGTGAACCCCGCTCTCTATAAAATACAAAAATTTGG	5340

QY		5341	CCGGGCAATGTGTGGCGGCACCTGTATAGTTCCAGCTACTCAGGAGGCTGAAGCAGAGAAATG	5400
Db		5341	CCGGGCAATGTGTGGCGGCACCTGTATAGTTCCAGCTACTCAGGAGGCTGAAGCAGAGAAATG	5400
QY		5401	GCGTGAACCCGAGAGGGCAGAGTTTGTCAGTGTAGGCCGAGATGCGCGCACTGCACATCCACACT	5460
Db		5401	GCGTGAACCCGAGAGGGCAGAGTTTGTCAGTGTAGGCCGAGATGCGCGCACTGCACATCCACACT	5460
QY		5461	GGGCGACAGACCAAGACTCCATCTGGAATAAAGAAAAGAAAGCTTCAGAGTCTGAGCCAGA	5520
Db		5461	GGGCGACAGACCAAGACTCCATCTGGAATAAAGAAAAGAAAGCTTCAGAGTCTGAGCCAGA	5520
QY		5521	GGCCCCAGGCTGTAAATTCTGTCACTTAACATGACCTTGGGCAAGGCACTTCCTTCCTCGGC	5580
Db		5521	GGCCCCAGGCTGTAAATTCTGTCACTTAACATGACCTTGGGCAAGGCACTTCCTTCCTCGGC	5580
QY		5581	CCAGTTCAAGGGGTTGGAATGAGACTCCAAAGTCCCCTTCAGCATTTAACGCTGCATGGTTTC	5640
Db		5581	CCAGTTCAAGGGGTTGGAATGAGACTCCAAAGTCCCCTTCAGCATTTAACGCTGCATGGTTTC	5640
QY		5641	TAAATGATGAAGATGAGGGGCACTTTCCCTCTCACCACCCAGCCGCTGCACATTCAGAGT	5700
Db		5641	TAAATGATGAAGATGAGGGGCACTTTCCCTCTCACCACCCAGCCGCTGCACATTCAGAGT	5700
QY		5701	GAATGACCAAGGGAAGTCAAGCTGTGTCCAATCCCGACAGTTCCAAAGCCCTTGGGAGCCCTAC	5760
Db		5701	GAATGACCAAGGGAAGTCAAGCTGTGTCCAATCCCGACAGTTCCAAAGCCCTTGGGAGCCCTAC	5760
QY		5761	TGTCAAGGTCCTGCACAGAGAGGTGAAGGTAGGTAGGCCAATGCTCTCGAAGGCTTTG	5820
Db		5761	TGTCAAGGTCCTGCACAGAGAGGTGAAGGTAGGTAGGCCAATGCTCTCGAAGGCTTTG	5820
QY		5821	CCTCATTTGGGACAAGACATCCGGTTTTCTCGTGGGCTACCGGGATTTCTAGGGGCTTTAGC	5880
Db		5821	CCTCATTTGGGACAAGACATCCGGTTTTCTCGTGGGCTACCGGGATTTCTAGGGGCTTTAGC	5880
QY		5881	CGAATGAGTCAATGAGGGGGCGGGGGGTTTCTGGGGGAGTTCCAGCTAATCAACTTGGGA	5940
Db		5881	CGAATGAGTCAATGAGGGGGCGGGGGGTTTCTGGGGGAGTTCCAGCTAATCAACTTGGGA	5940
QY		5941	CAGGACACCCCTGGAACTTTCGATGTCCTATCCAACTGTGGGGTGGGCACAGACGCCAA	6000
Db		5941	CAGGACACCCCTGGAACTTTCGATGTCCTATCCAACTGTGGGGTGGGCACAGACGCCAA	6000
QY		6001	GACCCAAATGTCCTATATCAGAGTAGGGGGCTCAGAGGTCACCCAGACAGGACAGCTCCGG	6060
Db		6001	GACCCAAATGTCCTATATCAGAGTAGGGGGCTCAGAGGTCACCCAGACAGGACAGCTCCGG	6060
QY		6061	AGAGTTTGGGGGTAGGAATGGGAGCAACACAGGCTTCTTTTCTCTCTTAGAATTTGGG	6120
Db		6061	AGAGTTTGGGGGTAGGAATGGGAGCAACACAGGCTTCTTTTCTCTCTTAGAATTTGGG	6120
QY		6121	GGCTTTGGGGGACAGGCTTGAAGATCCCAAAGAGAGGGGCAAAAGACATCCCCCAACAG	6180
Db		6121	GGCTTTGGGGGACAGGCTTGAAGATCCCAAAGAGAGGGGCAAAAGACATCCCCCAACAG	6180
QY		6181	TCTGCGAAGAGGAGAGAGGGGAGACCCCGAGCTCAGAGCTCCCACTCCCAACAGGCGCT	6235
Db		6181	TCTGCGAAGAGGAGAGAGGGGAGACCCCGAGCTCAGAGCTCCCACTCCCAACAGGCGCT	6235
RESULT 2				
AX250679	LOCUS	AX250679	6235 bp	DNA linear PAT 05-OCT-2001
DEFINITION	Sequence 6 from Patent WO0168882.			
ACCESSION	AX250679			
VERSION	AX250679.1 GI:15984423			
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			

AUTHORS Ivanov, E.
TITLE Methods for homologous recombination
JOURNAL Patent: WO 0168882-A 6 20-SEP-2001;
TRANSMARKOTIC THERAPIES, INC. (US)
FEATURES Location/Qualifiers
source 1. 6235
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1586 a 1610 c 1708 g 1331 t
ORIGIN
Query Match 100.0%; Score 6235; DB 6; Length 6235;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCAGTTGAGACAGTATGTTCAAGACAGCCCTGGGAGCATAGGAGACTGTCTCTACG 60
DB 1 GATCAGTTGAGACAGTATGTTCAAGACAGCCCTGGGAGCATAGGAGACTGTCTCTACG 60
QY 61 AAAAATCAAAAATATATGCGCGGCGATGCTGGCTAGCTGTATCCCTGAACCTTGGG 120
DB 61 AAAAATCAAAAATATATGCGCGGCGATGCTGGCTAGCTGTATCCCTGAACCTTGGG 120
QY 121 ACATCAAGGCAAGTGGATCACTTGAGTGCAGAGTTTCAGACTAGCTGGCCACATGTGT 180
DB 121 ACATCAAGGCAAGTGGATCACTTGAGTGCAGAGTTTCAGACTAGCTGGCCACATGTGT 180
QY 121 ACATCAAGGCAAGTGGATCACTTGAGTGCAGAGTTTCAGACTAGCTGGCCACATGTGT 180
DB 121 ACATCAAGGCAAGTGGATCACTTGAGTGCAGAGTTTCAGACTAGCTGGCCACATGTGT 180
QY 181 GAAACCCATATCTCCACTTAAAAATATCAAAATATAGCCAGGCAATGGTGACAGCCACTGTA 240
DB 181 GAAACCCATATCTCCACTTAAAAATATCAAAATATAGCCAGGCAATGGTGACAGCCACTGTA 240
QY 241 ATCCGGGTACTCAGAGGAGCTGAGGAGAGAGATCACTTGAACCCAGGAGGCGAGGTTG 300
DB 241 ATCCGGGTACTCAGAGGAGCTGAGGAGAGAGATCACTTGAACCCAGGAGGCGAGGTTG 300
QY 241 ATCCGGGTACTCAGAGGAGCTGAGGAGAGAGATCACTTGAACCCAGGAGGCGAGGTTG 300
DB 241 ATCCGGGTACTCAGAGGAGCTGAGGAGAGAGATCACTTGAACCCAGGAGGCGAGGTTG 300
QY 301 CAGTGAGCTGAGATCACACACTGCACCTCAGCCCTGGGTGACAGAGCAAGACTCTATCTC 360
DB 301 CAGTGAGCTGAGATCACACACTGCACCTCAGCCCTGGGTGACAGAGCAAGACTCTATCTC 360
QY 361 AAAAAAAT 420
DB 361 AAAAAAAT 420
QY 421 CTACTCAGAGGAGCTGAGGAGAGATCACTTGAACCTGGGAGCTCAAGGCTATAGTGA 480
DB 421 CTACTCAGAGGAGCTGAGGAGAGATCACTTGAACCTGGGAGCTCAAGGCTATAGTGA 480
QY 481 GCCAAGATCATGCCACTACCTCCAGCTGGGCAACAGAGAGAGACCCCTGTCTTAAAAA 540
DB 481 GCCAAGATCATGCCACTACCTCCAGCTGGGCAACAGAGAGAGACCCCTGTCTTAAAAA 540
QY 541 AAT 600
DB 541 AAT 600
QY 601 ATGTAT 660
DB 601 ATGTAT 660
QY 661 ATGTAT 720
DB 661 ATGTAT 720
QY 721 ACAT 780
DB 721 ACAT 780
QY 781 AAAAAT 840
DB 781 AAAAAT 840
QY 841 GCCATCGTACGCGCTGTATATCCAGACTTGGAGGCTGAGTGGGAGATCACTTGAAG 900

DB 841 GCCATCGTACGCGCTGTATATCCAGACTTGGAGGCTGAGTGGGAGATCACTTGAAG 900
QY 901 ATCAACAGTTGAGACAGCAGCTGGCCACATATGCAAAACCCCATTTCTACTTAAAAATATA 960
DB 901 ATCAACAGTTGAGACAGCAGCTGGCCACATATGCAAAACCCCATTTCTACTTAAAAATATA 960
QY 961 AAAAAT 1020
DB 961 AAAAAT 1020
QY 1021 GGAGATGCGCTTGAACCTGGAGAGGAGGTTGACAGTGGAGCCACATCTGACCTGCAC 1080
DB 1021 GGAGATGCGCTTGAACCTGGAGAGGAGGTTGACAGTGGAGCCACATCTGACCTGCAC 1080
QY 1081 TCCAGCTGGGTGACAGAGTGAAGTGTCTCTCAACGAAAAAAGTGTAAAGCCAT 1140
DB 1081 TCCAGCTGGGTGACAGAGTGAAGTGTCTCTCTCAACGAAAAAAGTGTAAAGCCAT 1140
QY 1141 TCCATATATGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 TCCATATATGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 ACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 ACCAAGTGGCCATCCAGAGAGGTTTTTTTTTACAATCTACACTCCCCCAGCAACAAAT 1320
DB 1261 ACCAAGTGGCCATCCAGAGAGGTTTTTTTTTACAATCTACACTCCCCCAGCAACAAAT 1320
QY 1321 GAGAGTACTCCAGATCTTTTCAAAAGATGCTTAAAGCCAGTACAGATGAAAAACAGA 1380
DB 1321 GAGAGTACTCCAGATCTTTTCAAAAGATGCTTAAAGCCAGTACAGATGAAAAACAGA 1380
QY 1381 AGTGAGAGGAGAGTGCAGGCGCTTCAACCATGAAAGAAATACCTGTTAGAGCCTTCT 1440
DB 1381 AGTGAGAGGAGAGTGCAGGCGCTTCAACCATGAAAGAAATACCTGTTAGAGCCTTCT 1440
QY 1441 GATGCTGGAAGAGATGAATACGGGGGTCTCTGAGAGCTGCGCCCTGTATGATCACTGTG 1500
DB 1441 GATGCTGGAAGAGATGAATACGGGGGTCTCTGAGAGCTGCGCCCTGTATGATCACTGTG 1500
QY 1501 ACTTGCAGACCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1560
DB 1501 ACTTGCAGACCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1560
QY 1561 CACTCTGTTTGGTCTTATATCTCCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTAT 1620
DB 1561 CACTCTGTTTGGTCTTATATCTCCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTAT 1620
QY 1621 TCAAGATGACAGCTTCTCTGACAGAGAAAGTGTCTACAGAAACAGCAGGGGCTTGGCA 1680
DB 1621 TCAAGATGACAGCTTCTCTGACAGAGAAAGTGTCTACAGAAACAGCAGGGGCTTGGCA 1680
QY 1681 AGAT 1740
DB 1681 AGAT 1740
QY 1741 AAGTTTTTCACTTCTGAGGCGCATGCCCTTGGCTACAAACACAGCAGTGTGTTGACAGA 1800
DB 1741 AAGTTTTTCACTTCTGAGGCGCATGCCCTTGGCTACAAACACAGCAGTGTGTTGACAGA 1800
QY 1801 TGAATATGAGAGTCCCTTACACCTGTATATCCACACACTTGGGAGGCGCAAGCGGGTGG 1860
DB 1801 TGAATATGAGAGTCCCTTACACCTGTATATCCACACACTTGGGAGGCGCAAGCGGGTGG 1860
QY 1861 ATGCTTGAAGCTGAGAGTACACAGATGCGGAGTCTCTACAGCCCTGCTTCCCTCTC 1920
DB 1861 ATGCTTGAAGCTGAGAGTACACAGATGCGGAGTCTCTACAGCCCTGCTTCCCTCTC 1920
QY 1921 GCGGCTCTCTGCTGCGGCTCCCACTTGGGAGGACTTGAAGAGCCCTTCAAGCCACCG 1980

Db 1921 GGGGCTCTCTGCTGGGCTCCACATTCCGCTGGACATTGAGAGGCCCTTACGCCACCG 1980
QY 1981 CTGACACTGTGGAGACCCCTTTCTGGGCTGGCCAAAGCCAGAGCCGGCTCCTCAGCTTGC 2040
Db 1981 CTGACACTGTGGAGACCCCTTTCTGGGCTGGCCAAAGCCAGAGCCGGCTCCTCAGCTTGC 2040
QY 2041 AGGAGAGTGTGAGGAGAGAGGCTCAGAGAGAACCGGGGCTGGCAGCGGGCTTGGGGG 2100
Db 2041 AGGAGAGTGTGAGGAGAGAGGCTCAGAGAGAACCGGGGCTGGCAGCGGGCTTGGGGG 2100
QY 2101 CAGCTGAGATTCCGGGTGGGGCTTGGCGGGGCCCGCAGTCGAGAGCAGCGGGCAG 2160
Db 2101 CAGCTGAGATTCCGGGTGGGGCTTGGCGGGGCCCGCAGTCGAGAGCAGCGGGCAG 2160
QY 2161 CCCCTGCCAGGCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGACGGGCTGGGAGGGGT 2220
Db 2161 CCCCTGCCAGGCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGACGGGCTGGGAGGGGT 2220
QY 2221 ACTGGGTGCCACAGTGGCCAGCCCGCGGCGCTGTGCTGGCTGATTTCTCACTGGGC 2280
Db 2221 ACTGGGTGCCACAGTGGCCAGCCCGCGGCGCTGTGCTGGCTGATTTCTCACTGGGC 2280
QY 2281 CTTAGCAGCCTTCCCGGGGAGAGGCTCGGGACCTGAGCCGCAATGCCCTGAGCCTCC 2340
Db 2281 CTTAGCAGCCTTCCCGGGGAGAGGCTCGGGACCTGAGCCGCAATGCCCTGAGCCTCC 2340
QY 2341 CTTTCATAGGGGCTCTGTGCGGCCCGGAGCCTTCCCGAGAGCAGCACCCCTGCTCCAG 2400
Db 2341 CTTTCATAGGGGCTCTGTGCGGCCCGGAGCCTTCCCGAGAGCAGCACCCCTGCTCCAG 2400
QY 2401 CGCCAGTCCCATGAGCAGCAGAGGGCTGAGAGTGGGGGAGCGACCGCGGAGCTGG 2460
Db 2401 CGCCAGTCCCATGAGCAGCAGAGGGCTGAGAGTGGGGGAGCGACCGCGGAGCTGG 2460
QY 2461 CAGGACAGTACCCCTGACGCCCTGTGTGCGGATCCAGTGGTGAAGCAGCTGGGCTCT 2520
Db 2461 CAGGACAGTACCCCTGACGCCCTGTGTGCGGATCCAGTGGTGAAGCAGCTGGGCTCT 2520
QY 2521 GAGTGTGGTGAAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATTGTAATACCAAT 2580
Db 2521 GAGTGTGGTGAAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATTGTAATACCAAT 2580
QY 2581 CAGCACCCTGTGCTAGCTCAGGGCTGTGTAATGCACCAATCCACACTCTGTATCTAGCT 2640
Db 2581 CAGCACCCTGTGCTAGCTCAGGGCTGTGTAATGCACCAATCCACACTCTGTATCTAGCT 2640
QY 2641 ACTCTGATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATTGTAATACCAATC 2700
Db 2641 ACTCTGATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATTGTAATACCAATC 2700
QY 2701 GGCACCTGTATCTAGCTCAAGGTTTAAACACCAATCAGCACCCTGTGTCTAGCTC 2760
Db 2701 GGCACCTGTATCTAGCTCAAGGTTTAAACACCAATCAGCACCCTGTGTCTAGCTC 2760
QY 2761 AGGATATGTAATGCACCAATTCGACACTGTGTATCTGCTTCAATGAGCAATCCGCT 2820
Db 2761 AGGATATGTAATGCACCAATTCGACACTGTGTATCTGCTTCAATGAGCAATCCGCT 2820
QY 2821 GAAGAGACCAACAGGCTTTGTGTGACCAATTAAGCTTATCAGCTGGGTGACGT 2880
Db 2821 GAAGAGACCAACAGGCTTTGTGTGACCAATTAAGCTTATCAGCTGGGTGACGT 2880
QY 2881 GGGCTGAGTCCGAAAAAGAGTCAAGGAGAGATTAAGGTTGGGGCTTTTATAGAT 2940
Db 2881 GGGCTGAGTCCGAAAAAGAGTCAAGGAGAGATTAAGGTTGGGGCTTTTATAGAT 2940
QY 2941 TTGGGTAGTAAAGAAAATTACAGTCAAGGGGGTTTGTCTCTGGCGGGCAGGAGTGG 3000
Db 2941 TTGGGTAGTAAAGAAAATTACAGTCAAGGGGGTTTGTCTCTGGCGGGCAGGAGTGG 3000
QY 3001 GGGGTGCAAGGTGCTCAGTGGGGGTCTTTTGAAGCAGAGTGAAGCAAGAAAGACT 3060
Db 3001 GGGGTGCAAGGTGCTCAGTGGGGGTCTTTTGAAGCAGAGTGAAGCAAGAAAGACT 3060

QY 3061 TTCAACAGGTAATGTCATCAATTAAAGCAGAGCCGCCAATTACACCTCTTTGTGGTG 3120
Db 3061 TTCAACAGGTAATGTCATCAATTAAAGCAGAGCCGCCAATTACACCTCTTTGTGGTG 3120
QY 3121 GAATGTCACTAGTTAAGTTGGGCGAGGGCAATTCACCTCTTTGTGTGATCTTCAAGTAC 3180
Db 3121 GAATGTCACTAGTTAAGTTGGGCGAGGGCAATTCACCTCTTTGTGTGATCTTCAAGTAC 3180
QY 3181 TTCAAGCCATCTGGCGCTATATGTGCAGATTACAGGGGATCCGATGGCTTGGTGGCT 3240
Db 3181 TTCAAGCCATCTGGCGCTATATGTGTGCAAGTTACAGGGGATCCGATGGCTTGGTGGCT 3240
QY 3241 CAGAGCTTGCACAGTACTGTGTGGGCTTTGGAGAAATGTTTGTGTGACACTCTGTAT 3300
Db 3241 CAGAGCTTGCACAGTACTGTGTGGGCTTTGGAGAAATGTTTGTGTGACACTCTGTAT 3300
QY 3301 CTAGTTAATCTAAGTGGGAGCTGGAGAACCTTTGTGTCTACTCAGGGATTGTAACCA 3360
Db 3301 CTAGTTAATCTAAGTGGGAGCTGGAGAACCTTTGTGTCTACTCAGGGATTGTAACCA 3360
QY 3361 CCAATCAGCGCCCTGTCAAAACAGACCACTGGCTCTACCAATCAGCAGATGTGGTGG 3420
Db 3361 CCAATCAGCGCCCTGTCAAAACAGACCACTGGCTCTACCAATCAGCAGATGTGGTGG 3420
QY 3421 GGCAGATTAAGAAATTAAGACAGGCTGCCGAGCCAGCAGTGGCAACGCCACAGGTCC 3480
Db 3421 GGCAGATTAAGAAATTAAGAAACAGAGCTGCCGAGCCAGCAGTGGCAACGCCACAGGTCC 3480
QY 3481 CTATCCAAATATGAGCAGCTTGTCTTTTCTGTCTTTGGCAATTAATCTGTCTACTGCTG 3540
Db 3481 CTATCCAAATATGAGCAGCTTGTCTTTTCTGTCTTTGGCAATTAATCTGTCTACTGCTG 3540
QY 3541 CTTTGTGGTCCACACTGCTTTTATGAGCTGTAACTACCTACCAAGAAAGTCTGACGCTC 3600
Db 3541 CTTTGTGGTCCACACTGCTTTTATGAGCTGTAACTACCTACCAAGAAAGTCTGACGCTC 3600
QY 3601 ACTCTGAGGCTAATGAGCAGAGCCACCGGAGAGATGAACAACTCCGGCGGCT 3660
Db 3601 ACTCTGAGGCTAATGAGCAGAGCCACCGGAGAGATGAACAACTCCGGCGGCT 3660
QY 3661 GCGTTAAGAGCTAATACACTCAGCGGAGGCTGCGAGCTTCACTCCAGCCAGCGAGA 3720
Db 3661 GCGTTAAGAGCTAATACACTCAGCGGAGGCTGCGAGCTTCACTCCAGCCAGCGAGA 3720
QY 3721 CCACGAACCCACGAAAGAAATGCGAACCATCTGAACTGAACTCAGAAAGAAACAACT 3780
Db 3721 CCACGAACCCACGAAAGAAATGCGAACCATCTGAACTGAACTCAGAAAGAAACAACT 3780
QY 3781 CCAAGTCAACCACTTAAGAGCTGTAAACTCACTGCGAGGGTCCGGCTCTCTTCTTG 3840
Db 3781 CCAAGTCAACCACTTAAGAGCTGTAAACTCACTGCGAGGGTCCGGCTCTCTTCTTG 3840
QY 3841 AAGTCACTGAGAACCAACACATCAACAGTTTCGACACAAAGCCGAGGTTTGAATAGC 3900
Db 3841 AAGTCACTGAGAACCAACACATCAACAGTTTCGACACAAAGCCGAGGTTTGAATAGC 3900
QY 3901 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTCAAAAAATTTGGCGG 3960
Db 3901 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTCAAAAAATTTGGCGG 3960
QY 3961 AGCATGTGTGCTCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Db 3961 AGCATGTGTGCTCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
QY 4021 TGAAGCTGGAGGAGGAGACTGCAAGTGAATGATTTGTCACACAGCCCTCTAGGCTGGG 4080
Db 4021 TGAAGCTGGAGGAGGAGACTGCAAGTGAATGATTTGTCACACAGCCCTCTAGGCTGGG 4080
QY 4081 GGAAGACTGAGACCTGTTTCCCTTCCGAAAAAAATTTGACAAAAATGTAAATAGAGT 4140
Db 4081 GGAAGACTGAGACCTGTTTCCCTTCCGAAAAAAATTTGACAAAAATGTAAATAGAGT 4140

4141 GCCTGATATGCTAGCGCAGTGGCTCATGCTGTATATCCAGACTTTGGGAACCGAG 4200
4141 GCCTGATATGCTAGCGCAGTGGCTCATGCTGTATATCCAGACTTTGGGAACCGAG 4200
4201 GCGGCGGGTCACTAAGGTCAGAGTGTAGACCAAGCTGGCAACATGAGAAAGCC 4260
4201 GCGGCGGGTCACTAAGGTCAGAGTGTAGACCAAGCTGGCAACATGAGAAAGCC 4260
4261 ATCTCTTAAAAAATACAAATTAACCGCTGTGGGGCAGTGTGGAGCATGCTGTAA 4320
4261 ATCTCTTAAAAAATACAAATTAACCGCTGTGGGGCAGTGTGGAGCATGCTGTAA 4320
4321 TCCGAGTACTAGAGGCTGAGGAGAGATCACTTGAACCCAGAGGCGGCTTGC 4380
4321 TCCGAGTACTAGAGGCTGAGGAGAGATCACTTGAACCCAGAGGCGGCTTGC 4380
4381 AGTGAAGGAGATGCTGCCATTTGCACTCCACCCTCCAGCTGGGCAACAGAGCCAA 4440
4441 CTCTGCTTAAAAAATACAAATTAACCGCTGTGGGGCAGTGTGGAGCATGCTGTAA 4500
4441 CTCTGCTTAAAAAATACAAATTAACCGCTGTGGGGCAGTGTGGAGCATGCTGTAA 4500
4501 TTGCGAGCAACATGTTTAAAGATGTGAGAGTCTGCTTCCATGCTCTGTAAAAAC 4560
4501 TTGCGAGCAACATGTTTAAAGATGTGAGAGTCTGCTTCCATGCTCTGTAAAAAC 4560
4551 CACCTTCAAGGCGAGGTCAGTGGCTCATGCTTATATCCAGACTTTGGAGGCGGAG 4620
4551 CACCTTCAAGGCGAGGTCAGTGGCTCATGCTTATATCCAGACTTTGGAGGCGGAG 4620
4621 GCGGGTGAATCACTGAGGTCAGAGTGTGAGAGCAGCTGACCAACATGAGTGAAT 4680
4621 GCGGGTGAATCACTGAGGTCAGAGTGTGAGAGCAGCTGACCAACATGAGTGAAT 4680
4681 CCCACCTTACTAAAAATACAAATTAAGTATGAGTGTGTCATGCTGTAAATCCAC 4740
4681 CCCACCTTACTAAAAATACAAATTAAGTATGAGTGTGTCATGCTGTAAATCCAC 4740
4741 CTACTTGGAGGCTAGGAGCAAGAAATCACTAAGAACAGGAGGCGGAGTGTAGTAG 4800
4741 CTACTTGGAGGCTAGGAGCAAGAAATCACTAAGAACAGGAGGCGGAGTGTAGTAG 4800
4801 CCGAGATGCTGCTTGAATGCTCAGCTGAGCAATGAGGAGAACTGCATCAAAAAAAC 4860
4801 CCGAGATGCTGCTTGAATGCTCAGCTGAGCAATGAGGAGAACTGCATCAAAAAAAC 4860
4861 AACCAAAAAAACCCTCTCTACTTCCAGGAGTGTGATACAGAGCTGGGCCATCACT 4920
4861 AACCAAAAAAACCCTCTCTACTTCCAGGAGTGTGATACAGAGCTGGGCCATCACT 4920
4921 GCAGGTCCTGAGCCACAGAGTAAAGCGGAGCTCAGAGACCGGAGCAAGATTAACAGT 4980
4921 GCAGGTCCTGAGCCACAGAGTAAAGCGGAGCTCAGAGACCGGAGCAAGATTAACAGT 4980
4981 TGTGATCAGTGTGATGATGAGAGTCCCTGCAATGTGTGATGATGATGATGATGAT 5040
4981 TGTGATCAGTGTGATGATGAGAGTCCCTGCAATGTGTGATGATGATGATGATGAT 5040
5041 AGCAGCAGAGATGAGCCATCCAGTACACATCCATCTCTCATCCAGAGATGCTGTT 5100
5041 AGCAGCAGAGATGAGCCATCCAGTACACATCCATCTCTCATCCAGAGATGCTGTT 5100
5101 CTGGCAGGCTGGGTAATTAAGACAGAGTGAACAGTCTTGGTGTGATGATGATGAT 5160
5101 CTGGCAGGCTGGGTAATTAAGACAGAGTGAACAGTCTTGGTGTGATGATGATGAT 5160
5161 TGCCCGAGGAGGCTTGGGCTGTAGAAAAAGTTCAGGCTTGGGCGGAGGAGCTGCG 5220
5161 TGCCCGAGGAGGCTTGGGCTGTAGAAAAAGTTCAGGCTTGGGCGGAGGAGCTGCG 5220
5221 TCACGCTGTATATCCAGCACTTTGGAGGCGGAGGCTGATCAGAGGTCAAGAGA 5280

5221 TCACGCTGTATATCCAGCACTTTGGAGGCGGAGGCTGATCAGAGGTCAAGAGA 5280
5281 TGTGACCATCTTGGCTAAACAGGTGAACCCCGCTCTACTTAAAAATACAAAAATGG 5340
5281 TGTGACCATCTTGGCTAAACAGGTGAACCCCGCTCTACTTAAAAATACAAAAATGG 5340
5341 CGGGCAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATGATG 5400
5341 CGGGCAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATGATG 5400
5401 GCGTAAACCCGAGAGGCTGAGAGTTCAGTGAACCCAGATGCGCCACTGATCCAGCT 5460
5401 GCGTAAACCCGAGAGGCTGAGAGTTCAGTGAACCCAGATGCGCCACTGATCCAGCT 5460
5461 GGGCGACAGAGCAAGTCTCATCTGAGAAAAAAGAAAAAAGCTTACGCTGAGCCAGA 5520
5461 GGGCGACAGAGCAAGTCTCATCTGAGAAAAAAGAAAAAAGCTTACGCTGAGCCAGA 5520
5521 GGGCGAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATGATG 5580
5521 GGGCGAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATGATG 5580
5581 CCAGTTCACGAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATG 5640
5581 CCAGTTCACGAGGCTGTATATCTGCTCACTTACATGATGATGATGATGATGATGATG 5640
5641 TAAGATGAGAGATGAGGCTGATTTCCCTCTCTCAACCCAGCCGCTGCTCAATCAAGT 5700
5641 TAAGATGAGAGATGAGGCTGATTTCCCTCTCTCAACCCAGCCGCTGCTCAATCAAGT 5700
5701 GAATGACAGGAGGAGTCACTGATCCCAATCCCGAGTTCCAAGGCTTGGGAGCCCTAC 5760
5701 GAATGACAGGAGGAGTCACTGATCCCAATCCCGAGTTCCAAGGCTTGGGAGCCCTAC 5760
5761 TGTCAAGGCTGTGACAGAGAGTGAAGTCAAGTGAAGCAATTCGCTGAGAGGCTTGG 5820
5761 TGTCAAGGCTGTGACAGAGAGTGAAGTCAAGTGAAGCAATTCGCTGAGAGGCTTGG 5820
5821 CCTCATTCGGGACAGATCCCGGTTCCCTCTGCTGCTACCGGAGTCTTGGGCTTAC 5880
5821 CCTCATTCGGGACAGATCCCGGTTCCCTCTGCTGCTACCGGAGTCTTGGGCTTAC 5880
5881 CGAATGAGTATGAGGAGGAGGAGGAGGAGTCTGAGGAGTTCAGTCAATCACTTGGGA 5940
5881 CGAATGAGTATGAGGAGGAGGAGGAGGAGTCTGAGGAGTTCAGTCAATCACTTGGGA 5940
5941 CAGGACAGGCTGGAATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 6000
5941 CAGGACAGGCTGGAATCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 6000
6001 GACCCATGCTCTTATCTCAGTGAAGGCTCAGAGGCTCCAGACAGGAGCTCCGG 6060
6001 GACCCATGCTCTTATCTCAGTGAAGGCTCAGAGGCTCCAGACAGGAGCTCCGG 6060
6061 AGAGTGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6120
6061 AGAGTGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6120
6121 GCGTGGGAGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6180
6121 GCGTGGGAGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6180
6181 TGTGACAG 6235
6181 TGTGACAG 6235

RESULT 3
ARI56461 6679 bp DNA linear PAT 08-AUG-2001
LOCUS ARI56461
DEFINITION Sequence 1 from patent US 6242218.
ACCESSION ARI56461

VERSION AR156461.1 GI:15125165
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6679)
AUTHORS Treco,D.A., Heartlein,M.W. and Selden,R.F.
TITLE Genomic sequences for protein production and delivery
JOURNAL Patent: US 6242218-A 1 05-JUN-2001;
FEATURES
source Location/Qualifiers
1..6679
/organism="unknown"
BASE COUNT 1684 a 1739 c 1830 g 1426 t
ORIGIN
Query Match 100.0%; Score 6235; DB 6; Length 6679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCAGCTTGAGGACAGTATTCAGACACCGCTGGGACGATAGGGAGACTGTCTACG 60
Db |
QY 20 GATCAGCTTGAGGACAGTATTCAGACACCGCTGGGACGATAGGGAGACTGTCTACG 79
Db |
QY 61 AAAAATCAAAAAATTATGGCCGGGATGTGGCTCAGCTGTATATCCCTGAACCTTTGGG 120
Db |
QY 80 AAAAATCAAAAAATTATGGCCGGGATGTGGCTCAGCTGTATATCCCTGAACCTTTGGG 139
Db |
QY 121 ACATCAAGGCAAGTGGATCAGCTTGAGGTCAGAGAGTTCAGAGCTAGCCTGGCCAACTGGT 180
Db |
QY 140 ACATCAAGGCAAGTGGATCAGCTTGAGGTCAGAGAGTTCAGAGCTAGCCTGGCCAACTGGT 199
Db |
QY 181 GAAACCCATCTCCACATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 240
Db |
QY 200 GAAACCCATCTCCACATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 259
Db |
QY 241 ATCCCGGCTACTCAGAGGCTGAGGACGAGAAATCACTTGAACCCAGAGGCGGAGGTTG 300
Db |
QY 260 ATCCCGGCTACTCAGAGGCTGAGGACGAGAAATCACTTGAACCCAGAGGCGGAGGTTG 319
Db |
QY 301 CAGTGAAGTATGATCACACCACTGCTCAGCCTGGGTGACAGCAAGACTCTATCTC 360
Db |
QY 320 CAGTGAAGTATGATCACACCACTGCTCAGCCTGGGTGACAGCAAGACTCTATCTC 379
Db |
QY 361 AAAAAAATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 420
Db |
QY 380 AAAAAAATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 439
Db |
QY 421 CTACTCAGAGGCTGAGGAGGATCAGCTTGAACCTGGGCGAGTAAAGGCTACAGTGA 480
Db |
QY 440 CTACTCAGAGGCTGAGGAGGATCAGCTTGAACCTGGGCGAGTAAAGGCTACAGTGA 499
Db |
QY 481 GCCAAGATCATGCACTACCTCAGCCTGGGCAACAGAGAGAGACCCCTGCTTAATAAA 540
Db |
QY 500 GCCAAGATCATGCACTACCTCAGCCTGGGCAACAGAGAGAGACCCCTGCTTAATAAA 559
Db |
QY 541 AAAAAAATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 600
Db |
QY 560 AAAAAAATCAAAAAATTATGCCAGCATGTGGTGGGACGACCTGTGTA 619
Db |
QY 601 ATGTATATAGTTTGAACCAAGATCAGATAGTCAATTTTTTTAGGCTTGGGGCGGT 660
Db |
QY 620 ATGTATATAGTTTGAACCAAGATCAGATAGTCAATTTTTTTAGGCTTGGGGCGGT 679
Db |
QY 661 ATGGTCTGTGCAAAATCACTGTGCCCTGTCTTTAGACAAAGAGAGCTATAAACAAT 720
Db |
QY 680 ATGGTCTGTGCAAAATCACTGTGCCCTGTCTTTAGACAAAGAGAGCTATAAACAAT 739
Db |
QY 721 ACATACATGAATTTTTTATAGACATGAGATTGGAATTTATATGATTTTAACTTTAT 780
Db |
QY 740 ACATACATGAATTTTTTATAGACATGAGATTGGAATTTATATGATTTTAACTTTAT 799
Db |
QY 781 AAAATATCTTTTAAAAATTTTCCCTTAACATTTTAAAAAGTAAAGCGGCGAGCGC 840
Db |

Db 800 AAAATATCTTTTAAAAATTTTCCCTTAACATTTTAAAAAGTAAAGCGGCGAGCGC 859
QY 841 GCCATCGTCAGCCCTGTAAATTCAGACATTTTGGAGGCTGAGTGGCATCTACTTGAG 900
Db |
QY 860 GCCATCGTCAGCCCTGTAAATTCAGACATTTTGGAGGCTGAGTGGCATCTACTTGAG 919
Db |
QY 901 ATCAACAGTTTGAACCAAGCTGGCCAAACATAGCAAAACCCATTTCTACTTAAATAAA 960
Db |
QY 920 ATCAACAGTTTGAACCAAGCTGGCCAAACATAGCAAAACCCATTTCTACTTAAATAAA 979
Db |
QY 961 AAAATTTAGCTGGGATGTGGTGCACACCTGTGATCCAGCTACTTGGGAGGCTGAGGCA 1020
Db |
QY 980 AAAATTTAGCTGGGATGTGGTGCACACCTGTGATCCAGCTACTTGGGAGGCTGAGGCA 1039
Db |
QY 1021 GGAGAAATCGCTTGAACCTGGGAAACGGAGGTTGGCATGAGCCAAATCATCTCAGTCAC 1080
Db |
QY 1040 GGAGAAATCGCTTGAACCTGGGAAACGGAGGTTGGCATGAGCCAAATCATCTCAGTCAC 1099
Db |
QY 1081 TCCAGCCTGGGTGACAGAGTGAACCTTGTCTCAACGAAAAAAAAGTGTAAAGCCAT 1140
Db |
QY 1100 TCCAGCCTGGGTGACAGAGTGAACCTTGTCTCAACGAAAAAAAAGTGTAAAGCCAT 1159
Db |
QY 1141 TCCTAATTCAGTGTACATCAGTGTACATCTCAGTGTGGGTACTCGGTCTGAGGCAT 1200
Db |
QY 1160 TCCTAATTCAGTGTACATCAGTGTACATCTCAGTGTGGGTACTCGGTCTGAGGCAT 1219
Db |
QY 1201 ACCTGAGAAAGTAGAGTGTGGTGTACAGAGCATACACATTTCCACATTAATAGACACT 1260
Db |
QY 1220 ACCTGAGAAAGTAGAGTGTGGTGTACAGAGCATACACATTTCCACATTAATAGACACT 1279
Db |
QY 1261 ACCAAGTTGCCATCCAGAGAGTGTGTGTGTACAAATCTACCTCCCCAGCAACAAT 1330
Db |
QY 1280 ACCAAGTTGCCATCCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1339
Db |
QY 1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCGATGATGATGAAAAAGGA 1380
Db |
QY 1340 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCGATGATGATGAAAAAGGA 1399
Db |
QY 1381 AGTGGAGGGGAAAGCTCCAGACCCCTTCTAACCATGAAAGAAATCTGTAGAGCCTTCT 1440
Db |
QY 1400 AGTGGAGGGGAAAGCTCCAGACCCCTTCTAACCATGAAAGAAATCTGTAGAGCCTTCT 1459
Db |
QY 1441 GGATGCTGGAAGATGAAATTAACGGGGGTCTGTGGAAGCTGCCCCCTGTCAATATCTGTG 1500
Db |
QY 1460 GGATGCTGGAAGATGAAATTAACGGGGGTCTGTGGAAGCTGCCCCCTGTCAATATCTGTG 1519
Db |
QY 1501 ACCTGAGGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1560
Db |
QY 1520 ACCTGAGGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1579
Db |
QY 1561 CACTCTGTGTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTAT 1620
Db |
QY 1580 CACTCTGTGTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTAT 1639
Db |
QY 1621 TCAAGATGTACAGCTTTCTTGACAGGAAAGTGTGTACAGAAACACAGAGGGGCTTGCA 1680
Db |
QY 1640 TCAAGATGTACAGCTTTCTTGACAGGAAAGTGTGTACAGAAACACAGAGGGGCTTGCA 1699
Db |
QY 1681 AGATGATCTTAAGTCAAAATCCTAGCTGGCTCAGGACACAGCTGTTGTGATCTTGAAC 1740
Db |
QY 1700 AGATGATCTTAAGTCAAAATCCTAGCTGGCTCAGGACACAGCTGTTGTGATCTTGAAC 1759
Db |
QY 1741 AAGTTTTTCACTTCTCTGAGGCAATCCCTTGGCTACACACACAGATTGGTTGACAGGA 1800
Db |
QY 1760 AAGTTTTTCACTTCTCTGAGGCAATCCCTTGGCTACACACACAGATTGGTTGACAGGA 1819
Db |
QY 1801 TGAATGACGAATGCTTTACACCTGTAAATCCAGCACTTTGGAGGCGCAAGCGGGTGG 1860
Db |
QY 1820 TGAATGACGAATGCTTTACACCTGTAAATCCAGCACTTTGGAGGCGCAAGCGGGTGG 1879
Db |
QY 1861 ATGGCTTGAGCGTGAGAGGTGACAGCATGGCGGCAAGTCTCAAGAGCCCTGTCGCTTC 1920
Db |
QY 1880 ATGGCTTGAGCGTGAGAGGTGACAGCATGGCGGCAAGTCTCAAGAGCCCTGTCGCTTC 1939
Db |

QY	1921	GGCGCTCTCTGCCCTGGGCTCCACATTCGGTGACATTTGAGGAGCCCTTCAGCCACCG	1980
Db	1940	GGCGCTCTCTCTGGCTGGGCTCCACATTCGGTGACATTTGAGGAGCCCTTCAGCCACCG	1999
QY	1981	CTGCACGTGGGAGCCCTTCTCTGGGCTGGGCAAGGCCAGAGCCGGCTCCCTCAGCTTGC	2040
Db	2000	CTGCACGTGGGAGCCCTTCTCTGGGCTGGGCAAGGCCAGAGCCGGCTCCCTCAGCTTGC	2059
QY	2041	AGGAGGTGTGGAGGAGAGGCTCAAGCAGAAACCGGGGCTGGCCAGCGGCTTTCGGGCG	2100
Db	2060	AGGAGGTGTGGAGGAGAGGCTCAAGCAGAAACCGGGGCTGGCCAGCGGCTTTCGGGCG	2119
QY	2101	CAGCTGGAGTTCGGGGTGGGCTGTGGGCTTGGCCGGCCCGCATCTCGAAGCAGCGCCAG	2160
Db	2120	CAGCTGGAGTTCGGGGTGGGCTGTGGGCTTGGCCGGCCCGCATCTCGAAGCAGCGCCAG	2179
QY	2161	CCCTGACAGGCCCCGGGCAATGAGAAGGCTTGACACCCGGGCGCAGCGGCTGGAGGGGT	2220
Db	2180	CCCTGACAGGCCCCGGGCAATGAGAAGGCTTGACACCCGGGCGCAGCGGCTGGAGGGGT	2239
QY	2221	ACTGGGTGCCCAAGCAGTGCACAGCCCGCGCGCTGTGCTCGTCGATTTCTCACTGGGC	2280
Db	2240	ACTGGGTGCCCAAGCAGTGCACAGCCCGCGCGCTGTGCTCGTCGATTTCTCACTGGGC	2299
QY	2281	CTTAGAGACCTTCCCGGGGGGCGAGGGCTCGGAGACCTGACAGCCGCGCATGGCTGAGCCTC	2340
Db	2300	CTTAGAGACCTTCCCGGGGGGCGAGGGCTCGGAGACCTGACAGCCGCGCATGGCTGAGCCTC	2359
QY	2341	CTTCATGGGCTCTGTGCGGGCCGAGCCCTCCGAGCAGCACCAACCCCTGCTCCACAG	2400
Db	2360	CTTCATGGGCTCTGTGCGGGCCGAGCCCTCCGAGCAGCACCAACCCCTGCTCCACAG	2419
QY	2401	GGCCAGTCCCATGAGCACCAGAGGGCTGAGAGTGGGGCGACAGGCCACCGGAGCTGG	2460
Db	2420	GGCCAGTCCCATGAGCACCAGAGGGCTGAGAGTGGGGCGACAGGCCACCGGAGCTGG	2479
QY	2461	CAGGACGTACCCCTGCGAGCCCTGGTGGGANTCCACTGAGGTGGAAGCAGCTGGGCTCT	2520
Db	2480	CAGGACGTACCCCTGCGAGCCCTGGTGGGANTCCACTGAGGTGGAAGCAGCTGGGCTCT	2539
QY	2521	GAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGAGTTCGTAATACACCAAT	2580
Db	2540	GAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGAGTTCGTAATACACCAAT	2599
QY	2581	CAGCACCTGTGTAGCTCAGGGCTGTGTGATCGACCAATCCACACTCTGTATCTAGCT	2640
Db	2600	CAGCACCTGTGTAGCTCAGGGCTGTGTGATCGACCAATCCACACTCTGTATCTAGCT	2659
QY	2641	ACTCTGATGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTTCGTAATACACCAATC	2700
Db	2660	ACTCTGATGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTTCGTAATACACCAATC	2719
QY	2701	GGCACTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGAACCTTGTGTAGCTC	2760
Db	2720	GGCACTGTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGAACCTTGTGTAGCTC	2779
QY	2761	AGGCTATGTGAATGCACCAATCGACAGTCTGTATCTGGCTACTTTCATGGGCATCCGTGT	2820
Db	2780	AGGCTATGTGAATGCACCAATCGACAGTCTGTATCTGGCTACTTTCATGGGCATCCGTGT	2839
QY	2821	GAAAGAACCCACCAACAGGCTTGTGTGAGCAATTAAGCTTCTATCACCTGGGTGACAGT	2880
Db	2840	GAAAGAACCCACCAACAGGCTTGTGTGAGCAATTAAGCTTCTATCACCTGGGTGACAGT	2899
QY	2881	GGGCTGATCCGAAAAGAGAGTCAAGCAAGGAGATAAGGCTGGGGCCGTTTATAGAT	2940
Db	2900	GGGCTGATCCGAAAAGAGAGTCAAGCAAGGAGATAAGGCTGGGGCCGTTTATAGAT	2959
QY	2941	TTGGGTAGGTAAAGAAAATTAACGTCAAGAGGGGTTTGTCTGTGGCGGCGAGAGTGG	3000
Db	2960	TTGGGTAGGTAAAGAAAATTAACGTCAAGAGGGGTTTGTCTGTGGCGGCGAGAGAGG	3019

QY	3001	GGGGTCGCAAGGTCCTCAGTGGGGGGTCTTTTTCAGCCAGATCAGCCAGGAAAGACT	3060
Db	3020	GGGGTCGCAAGGTGCTCAGTGGGGGGTCTTTTTCAGCCAGGATCAGCCAGGAAAGACT	3079
QY	3061	TTCCAAAGGTATGTCATCAATTTAAAGCAAGACCCCGCAATTACCTCTTTTGTGTGTG	3120
Db	3080	TTCCAAAGGTATGTCATCAATTTAAAGCAAGACCCCGCAATTACCTCTTTTGTGTGTG	3139
QY	3121	GAATGTCAATCAGTTAAGTTGGGGGACAGGGCAATTTCACCTCTTTTGTGATTCCTTCAGTTAC	3180
Db	3140	GAATGTCAATCAGTTAAGTTGGGGGACAGGGCAATTTCACCTCTTTTGTGATTCCTTCAGTTAC	3199
QY	3181	TTCAGGCCATCTGGGCGTATATATGTGCAAGTTACAGGGAGTCGATGGCTTGCTGGGCT	3240
Db	3200	TTTACGGCCATCTGGGCGTATATATGTGCAAGTTACAGGGAGTCGATGGCTTGCTGGGCT	3259
QY	3241	CAGAGCCTTGACAGCTACTCTGTGTGGGGGCTTTGGAGAAATGTTTGTGTGCACACTCTGAT	3300
Db	3260	CAGAGCCTTGACAGCTACTCTGTGTGGGGGCTTTGGAGAAATGTTTGTGTGCACACTCTGAT	3319
QY	3301	CTAGTTAATCTAGTGGGGACGTGGAGAACCTTTTGTGTCAGTCCAGGGATTTGTAAAGCA	3360
Db	3320	CTAGTTAATCTAGTGGGGACGTGGAGAACCTTTTGTGTCAGTCCAGGGATTTGTAAAGCA	3379
QY	3361	CCAAATCAGCGCCCTGTCTCAAAAACAGACACTCGGCTCTTACCAATCAGCAGAGATGTGGTGG	3420
Db	3380	CCAAATCAGCGCCCTGTCTCAAAAACAGACACTCGGCTCTTACCAATCAGCAGAGATGTGGTGG	3439
QY	3421	GGCCAGATTAAGATTAATAAACGAGGTGCCCGGACCCAGACAGTGGCAACCGGCACAGTCC	3480
Db	3440	GGCCAGATTAAGATTAATAAACGAGGTGCCCGGACCCAGACAGTGGCAACCGGCACAGTCC	3499
QY	3481	CTATCCACAATATGGCAGCTTTGTTCTTTTCTGTTCGATTAATCTTTCGACTACTGTGTCG	3540
Db	3500	CTATCCACAATATGGCAGCTTTGTTCTTTTCTGTTCGATTAATCTTTCGACTACTGTGTCG	3559
QY	3541	CTTTTGGGTCACACTGCTTTTATAGCTGTAACTACACAGGAAGTCTGCAGCTTC	3600
Db	3560	CTTTTGGGTCACACTGCTTTTATAGCTGTAACTACACTACACAGGAAGTCTGCAGCTTC	3619
QY	3601	ACTCTGAAGGCACATTAAGACAGAGCCACCGGAGGAATGAACAACTCCGGCGGGCT	3660
Db	3620	ACTCTGAAGGCACATTAAGACAGAGCCACCGGAGGAATGAACAACTCCGGCGGGCT	3679
QY	3661	GCCTTAAAGGTATTAACACTACCGCGAAGTCTGCAGCTTCACTCCTCAGCCAGCGAGA	3720
Db	3680	GCCTTAAAGGTATTAACACTACCGCGAAGTCTGCAGCTTCACTCCTCAGCCAGCGAGA	3739
QY	3721	CCACGAACCCACAGAGGAAGAACTGCGAACACATCTGACATCTCAGAGGAACAACT	3780
Db	3740	CCACGAACCCACAGAGGAAGAACTGCGAACACATCTGACATCTCAGAGGAACAACT	3799
QY	3781	CCAAGTGCACCACTTAAGACCTGTAAACACTCACTGCGAGGGTCCGGGCTTCTCTTGTG	3840
Db	3800	CCAAGTGCACCACTTAAGACCTGTAAACACTCACTGCGAGGGTCCGGGCTTCTCTTGTG	3859
QY	3841	AAGTAGTAGACCCAGACACTCACAGTTTGGACACACAAAGCCACAGAGTTTGAATCAGC	3900
Db	3860	AAGTAGTAGACCCAGACACTCACAGTTTGGACACACAAAGCCACAGAGTTTGAATCAGC	3919
QY	3901	CTGGGCAACATGATGAAATGCCCTCTCTGCAAAAAAATTAACAAAAATTTGGCGG	3960
Db	3920	CTGGGCAACATGATGAAATGCCCTCTCTGCAAAAAAATTAACAAAAATTTGGCGG	3979
QY	3961	AGCATGTGTGTCCTGTGCTGTGTGCCAGCTACGCGGGAGGCTAAAGTGGAGATGCT	4020
Db	3980	AGCATGTGTGTCCTGTGCTGTGTGCCAGCTACGCGGGAGGCTAAAGTGGAGATGCT	4039
QY	4021	TGAGCGCTGGGAGGTGAAGACGTGACAGTGAAGCTGTATTTGTACCAAGCCCTCTAGGCTGGG	4080
Db	4040	TGAGCGCTGGGAGGTGAAGACGTGACAGTGAAGCTGTATTTGTACCAAGCCCTCTAGGCTGGG	4099
QY	4081	GGACAGACTGAGACCTCTTTCCCTCCGCAAAAAAATTTGACAAAGCTGTAAATGAAGCT	4140

LOCUS AX250678 6679 bp DNA linear PAT 05-OCT-2001
 DEFINITION Sequence 5 from Patent WO0168882.
 ACCESSION AX250678
 VERSION AX250678.1 GI:15984422
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Ivanov, E.
 AUTHORS
 TITLE Methods for homologous recombination
 JOURNAL Patent: WO 0168882-A 5 20-SEP-2001;
 TRANSMUTIC THERAPIES, INC. (US)
 FEATURES
 source location/Qualifiers
 1. 6679
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 1684 a 1739 c 1830 g 1426 t
 ORIGIN
 Query Match 100.0%; Score 6235; DB 6; Length 6679;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GATCCTGAGGAGCTAGTTCACAGCCAGCTGGGAGCATAGGAGACTCTCTAG 60
 20 GATCCTGAGGAGCTAGTTCACAGCCAGCTGGGAGCATAGGAGACTCTCTAG 79
 61 AAAATCAAAAATATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 80 AAAATCAAAAATATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 139
 121 ACATCAGGCAAGTGGATCCTTGGAGGTCAGAGTTCAGACTAGCTGGCCACATG 180
 140 ACATCAGGCAAGTGGATCCTTGGAGGTCAGAGTTCAGACTAGCTGGCCACATG 199
 181 GAACCCCTATCTCCACTAAAAATTAACAGGCGATGGTGGAGGCGACCTGTA 240
 200 GAACCCCTATCTCCACTAAAAATTAACAGGCGATGGTGGAGGCGACCTGTA 259
 241 ATCCGGCTACTCAGAGGCTGGAGGAGAGATCACTTGAACCCAGAGGCGGAGTTG 300
 260 ATCCGGCTACTCAGAGGCTGGAGGAGAGATCACTTGAACCCAGAGGCGGAGTTG 319
 301 CAGTACGTGAGATCACACACTGCTCAGCTGGGTCAGAGCAAGCACTATCTC 360
 320 CAGTACGTGAGATCACACACTGCTCAGCTGGGTCAGAGCAAGCACTATCTC 379
 361 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
 380 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 439
 421 CTACTCAGAGGCTGAGTGGAGGATCACTTGAACCTGGGCACTCAAGCTAGAGTA 480
 440 CTACTCAGAGGCTGAGTGGAGGATCACTTGAACCTGGGCACTCAAGCTAGAGTA 499
 481 GCCAAGATCAGCTACACTCAGCTGGGCAAGAGAGAGAGAGAGAGAGAGAGAG 540
 500 GCCAAGATCAGCTACACTCAGCTGGGCAAGAGAGAGAGAGAGAGAGAGAGAG 559
 541 AATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 560 AATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 619
 601 ATGTATATAGTTTGAAGAACTCAAGATCAGATATGCAATTTTAAAGGCTTGGCCGT 660
 620 ATGTATATAGTTTGAAGAACTCAAGATCAGATATGCAATTTTAAAGGCTTGGCCGT 679
 661 ATGTCTCTGTACAACTACCTGCTCTTCTTCTAGACAAAGAGCAAGCTTAAACAT 720
 680 ATGTCTCTGTACAACTACCTGCTCTTCTTCTAGACAAAGAGCAAGCTTAAACAT 739

721 ACATACGAAATTTTATATAGACATGAGATTTGAATTTATATGATTTTACATTTAT 780
 740 ACATACGAAATTTTATATAGACATGAGATTTGAATTTATATGATTTTACATTTAT 799
 781 AAAATATCTTTTAAAAATTTTCCCTAACATTTAAAAAGTAAAGCCGGCCAGCC 840
 800 AAAATATCTTTTAAAAATTTTCCCTAACATTTAAAAAGTAAAGCCGGCCAGCC 859
 841 GCCATCTGACGCTGTATTTCCAGACTTTGGAGGCTGAGGTGGGAGATCACTTGA 900
 860 GCCATCTGACGCTGTATTTCCAGACTTTGGAGGCTGAGGTGGGAGATCACTTGA 919
 901 ATCAACAGTTGAGAGCCAGCTGGCCCAATATGCAAAACCCATTTCTCTAAATAA 960
 920 ATCAACAGTTGAGAGCCAGCTGGCCCAATATGCAAAACCCATTTCTCTAAATAA 979
 961 AAAATTAGCTGGGATAGTGTGACACCTGTATCCAGCTTGGAGGCTGAGGCA 1020
 980 AAAATTAGCTGGGATAGTGTGACACCTGTATCCAGCTTGGAGGCTGAGGCA 1039
 1021 GGAGAAATGCTTGAACCTGGGAAAGGAGGTTGACAGCAATCATGCCACTGCAC 1080
 1040 GGAGAAATGCTTGAACCTGGGAAAGGAGGTTGACAGCAATCATGCCACTGCAC 1099
 1081 TCCAGGCTGGGATGACAGTGGAGTGGTCTCAAGAAAAAAGTAAAGCCAT 1140
 1100 TCCAGGCTGGGATGACAGTGGAGTGGTCTCAAGAAAAAAGTAAAGCCAT 1159
 1141 TCCATATATGATACATAGTATCATAGTCTGCTGCTGCTGAGGCAAT 1200
 1160 TCCATATATGATACATAGTATCATAGTCTGCTGCTGCTGAGGCAAT 1219
 1201 ACCTGAGAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 1220 ACCTGAGAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1279
 1261 ACCAAGTGGCAATCCAGAGAGGTTTTTTTAAATCAATCACTGCTGCTGCTGCTG 1320
 1280 ACCAAGTGGCAATCCAGAGAGGTTTTTTTAAATCAATCACTGCTGCTGCTGCTG 1339
 1321 GAGAGTACTCCAGATCTTTCAAAAGATGCTTAAGCCAGTACAGATGAAGACAGA 1380
 1340 GAGAGTACTCCAGATCTTTCAAAAGATGCTTAAGCCAGTACAGATGAAGACAGA 1399
 1381 AGTGGAGGGAAGCTGCCAGCCCTTCAACATGTAAGAAATACCTGTAAGACCTTCT 1440
 1400 AGTGGAGGGAAGCTGCCAGCCCTTCAACATGTAAGAAATACCTGTAAGACCTTCT 1459
 1441 GGATGCTGGAAGATGAATTAAGGGGGTCTGAGGCTGGCCCTGCTGAGATCACTG 1500
 1460 GGATGCTGGAAGATGAATTAAGGGGGTCTGAGGCTGGCCCTGCTGAGATCACTG 1519
 1501 ACTTGAAGCTTCAGATCCAGTCCAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 1520 ACTTGAAGCTTCAGATCCAGTCCAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTG 1579
 1561 CACTCTGTTGGTCTTATTTCTCCCATGTTGGGCTGAGCTGAGTTGAGCCGTAT 1620
 1580 CACTCTGTTGGTCTTATTTCTCCCATGTTGGGCTGAGCTGAGTTGAGCCGTAT 1639
 1621 TCAAGATGTAAGCTTTCTTGAAGAGAAAGTATGTCACAGAAACAGAGGGGCTTGG 1680
 1640 TCAAGATGTAAGCTTTCTTGAAGAGAAAGTATGTCACAGAAACAGAGGGGCTTGG 1699
 1681 AGATGATCTAATGCAAAATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 1700 AGATGATCTAATGCAAAATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1759
 1741 AAGTTTTTCACTTCTGAGGCAATCCCTTGGCTAACAACAGAGTTGTTGACAGA 1800
 1760 AAGTTTTTCACTTCTGAGGCAATCCCTTGGCTAACAACAGAGTTGTTGACAGA 1819

QY 1801 TGAATACAGAGTCCCTTACACTGTATATCCAGCACTTTGGAGGCCAAAGCGGGGTGG 1860
Db 1820 TGAATACAGAGTCCCTTACACTGTATATCCAGCACTTTGGAGGCCAAAGCGGGGTGG 1879
QY 1861 ATGGCTTGAGAGCTGAGAGTGACAGATGCGGAGATCTCCACAGCCCTGGTTGCTCTC 1920
Db 1880 ATGGCTTGAGAGCTGAGAGTGACAGATGCGGAGATCTCCACAGCCCTGGTTGCTCTC 1939
QY 1921 GGGGCTCTCTCTGCTGGGCTCCCACTTCCGTGGCACTTGAGAGACCCCTTACGCCCAAG 1980
Db 1940 GGGGCTCTCTCTGCTGGGCTCCCACTTCCGTGGCACTTGAGAGACCCCTTACGCCCAAG 1999
QY 1981 CTGCACTGTGGAGACCCCTTTCTGGCTGGCCAAAGGCCAGAGCCGCTCTCAAGCTTGC 2040
Db 2000 CTGCACTGTGGAGACCCCTTTCTGGCTGGCCAAAGGCCAGAGCCGCTCTCTCAAGCTTGC 2059
QY 2041 AGGAGAGTGTGAGAGGAGAGGCTCAAGCAGAACCGGGGCTGGCCAGAGCCGCTTGGGGC 2100
Db 2060 AGGAGAGTGTGAGAGGAGAGGCTCAAGCAGAACCGGGGCTGGCCAGAGCCGCTTGGGGC 2119
QY 2101 CAGCTGAGTTCGCGGTGGCGGTGGCTTGGCGGGCCCGCACTCGAGAGCCGCGCAG 2160
Db 2120 CAGCTGAGTTCGCGGTGGCGGTGGCTTGGCGGGCCCGCACTCGAGAGCCGCGCAG 2179
QY 2161 CCCGCGCAGGGCCCCGGGCANTGAGAGGCTTACAGCACCCGGGCGCAGCGGCTGGAGAGGTGT 2220
Db 2180 CCCGCGCAGGGCCCCGGGCANTGAGAGGCTTACAGCACCCGGGCGCAGCGGCTGGAGAGGTGT 2239
QY 2221 ACTGGGTGCCCCAGCAGTGGCAGCCCGCGCGCTGTGCTGCTGATTTTCTCACTGGGC 2280
Db 2240 ACTGGGTGCCCCAGCAGTGGCAGCCCGCGCGCTGTGCTGCTGATTTTCTCACTGGGC 2299
QY 2281 CTTAGCAGCTTCCCCGGGGGCGAGGGCTCGGGACCTCGAGCCGCCATGCTGAGCCTCC 2340
Db 2300 CTTAGCAGCCTTCCCCGGGGGCGAGGGCTCGGGACCTCGAGCCGCCATGCTGAGCCTCC 2359
QY 2341 CCTCATAGGGGCTCCTGTGTGGCGCCGAGAGCTCCCGAGAGAGCAACCCCGCTGCCACAG 2400
Db 2360 CCTCATAGGGGCTCCTGTGTGGCGCCGAGAGCTCCCGAGAGAGCAACCCCGCTGCCACAG 2419
QY 2401 CGCCCACTCCCATGACACAGCAAGGGCTGAGAACTGGGGCGCACGGCAGCCGGAGCTGG 2460
Db 2420 CGCCCACTCCCATGACACAGCAAGGGCTGAGAACTGGGGCGCACGGCAGCCGGAGCTGG 2479
QY 2461 CAGGACACTACCCCTGCAAGCCCTGTGTGGGATCCACTGGGTGAAGCCAGCTGGGCTCT 2520
Db 2480 CAGGACACTACCCCTGCAAGCCCTGTGTGGGATCCACTGGGTGAAGCCAGCTGGGCTCT 2539
QY 2521 GAGCTGTGGAGACTTGGAGAACCTTATGTCTAGCTCAGGGATTCGTAATTAACCAAT 2580
Db 2540 GAGCTGTGGAGACTTGGAGAACCTTATGTCTAGCTCAGGGATTCGTAATTAACCAAT 2599
QY 2581 CAGCACCCTGTGTCTAGCTCAGGGTGTGTGAATGACCAATCCACACTCTGTATCTAGCT 2640
Db 2600 CAGCACCCTGTGTCTAGCTCAGGGTGTGTGAATGACCAATCCACACTCTGTATCTAGCT 2659
QY 2641 ACTGTGATGGGGCTTGGAGAACCTTATGTCTAGCTCAGGGATTCGTAATTAACCAATC 2700
Db 2660 ACTGTGATGGGGCTTGGAGAACCTTATGTCTAGCTCAGGGATTCGTAATTAACCAATC 2719
QY 2701 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTGTAGCTC 2760
Db 2720 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTGTAGCTC 2779
QY 2761 AGGCTATGTAAATGACCAATTCAGCAGTCTGTATCTGGTACTTCAATGGGCATCCGCT 2820
Db 2780 AGGCTATGTAAATGACCAATTCAGCAGTCTGTATCTGGTACTTCAATGGGCATCCGCT 2839
QY 2821 GAAGAGACCAACCAAGGCTTGTGTGAGCAATTAAGCTTCTATCACTGGGTGACAGT 2880
Db 2840 GAAGAGACCAACCAAGGCTTGTGTGAGCAATTAAGCTTCTATCACTGGGTGACAGT 2899
QY 2881 GGGCTGAGTCCGAAAGAGAGTCAAGCAAGGAGATTAAGGTGGGCGCTTTTATAGAT 2940

Db 2900 GGGCTGAGTCCGAAAGAGAGTCAAGCAGAGGAGATTAAGGTGGGCGCTTTTATAGAT 2959
QY 2941 TTGGTGTAGTAAAGAAAATTAACAGTCAAAAGGGGTTTGTCTGTGGGGGCGACAGAGG 3000
Db 2960 TTGGTGTAGTAAAGAAAATTAACAGTCAAAAGGGGTTTGTCTGTGGGGGCGACAGAGG 3019
QY 3001 GGGGTGCGAAGGTCTCAGTGGGGGTGCTTTTGTAGCCAGGATGAGCCAGGAAAAGACT 3060
Db 3020 GGGGTGCGAAGGTCTCAGTGGGGGTGCTTTTGTAGCCAGGATGAGCCAGGAAAAGACT 3079
QY 3061 TTCAAGGTAAATGTCAATTAAGGCAAGACCAGCATTTACACTCTTTTGTGTG 3120
Db 3080 TTCAAGGTAAATGTCAATTAAGGCAAGACCAGCATTTACACTCTTTTGTGTG 3139
QY 3121 GAATGTCAATCAGTTAAGTTGGGGAGGGCAATTCACCTTTTGTGTATTTCAATAC 3180
Db 3140 GAATGTCAATCAGTTAAGTTGGGGAGGGCAATTCACCTTTTGTGTATTTCAATAC 3199
QY 3181 TTCAAGCCATCTGGGCGTATATGTCAAGTTACAGGGGATGCAATGGCTTGGCT 3240
Db 3200 TTCAAGCCATCTGGGCGTATATGTCAAGTTACAGGGGATGCAATGGCTTGGCT 3259
QY 3241 CAGAGGCTTGACAGTACTCTGTGTGGGCTTTGGAGAAATTTTGTGTGACACTGTAT 3300
Db 3260 CAGAGGCTTGACAGTACTCTGTGTGGGCTTTGGAGAAATTTTGTGTGACACTGTAT 3319
QY 3301 CTAGTTAATCTAGTGGGAGCTGAGAACTTTTGTGTACTCAGGATTTGAACGA 3360
Db 3320 CTAGTTAATCTAGTGGGAGCTGAGAACTTTTGTGTACTCAGGATTTGAACGA 3379
QY 3361 CCAATCAGCGCCCTGTCAAAAACAGACCTCGGCTTACCAATCAGCAGATGTGGGTGG 3420
Db 3380 CCAATCAGCGCCCTGTCAAAAACAGACCTCGGCTTACCAATCAGCAGATGTGGGTGG 3439
QY 3421 GGCAGATTAAGAAATTAACAGAGCTGCCGACCGAGCAAGTGGCAACGCCACAGATCC 3480
Db 3440 GGCAGATTAAGAAATTAACAGAGCTGCCGACCGAGCAAGTGGCAACGCCACAGATCC 3499
QY 3481 CTATCCACAAATATGGAGCTTTGCTTTTGTGTGTGCAATATCTGTACTGTCTG 3540
Db 3500 CTATCCACAAATATGGAGCTTTGCTTTTGTGTGTGCAATATCTGTACTGTCTG 3559
QY 3541 CTTTTTGGGTCCACACTGCTTTTATAGCTGTAACTCACCAGAAAGTCTGCACTTC 3600
Db 3560 CTTTTTGGGTCCACACTGCTTTTATAGCTGTAACTCACCAGAAAGTCTGCACTTC 3619
QY 3601 ACTCCTGAAGCCACTAAGACACAGAGCCACCGGAGGAGATGAACAACTCCGGCGGCT 3660
Db 3620 ACTCCTGAAGCCACTAAGACACAGAGCCACCGGAGGAGATGAACAACTCCGGCGGCT 3679
QY 3661 GCCTTAAGAGCTATTAACACTCAGCGGAAAGGTGTGCACTTCACTCCACAGCGAGA 3720
Db 3680 GCCTTAAGAGCTATTAACACTCAGCGGAAAGGTGTGCACTTCACTCCACAGCGAGA 3739
QY 3721 CCACGAACCCACAGAAAGAAACTGCGAACACATCTGAATCAGAAAGAACAACT 3780
Db 3740 CCACGAACCCACAGAAAGAAACTGCGAACACATCTGAATCAGAAAGAACAACT 3799
QY 3781 CCAATGACACACTTAAAGAGCTTAACTCAGTGGAGGGTCCGGGCTTCTTCTTG 3840
Db 3800 CCAATGACACACTTAAAGAGCTTAACTCAGTGGAGGGTCCGGGCTTCTTCTTG 3859
QY 3841 AAGTCAGTGAAGCAACAGTCAACAGTTTGGACACAAGCCCGAGGTTTGAATACG 3900
Db 3860 AAGTCAGTGAAGCAACAGTCAACAGTTTGGACACAAGCCCGAGGTTTGAATACG 3919
QY 3901 CTGGGCAACATGATGAATGCCCTCTGTGCAAAAAAATTAACAAAAATTTGGCGG 3960
Db 3920 CTGGGCAACATGATGAATGCCCTCTGTGCAAAAAAATTAACAAAAATTTGGCGG 3979
QY 3961 AGCATGTGTGCTGCTGTGTGCTCCAGCTACCGGGAGGCTTAAAGTGGAGATGCT 4020

Db 3980 AGCATGGTGGTCCGTGCTGTCGCCAGCTACGGGGAGGCTAAAGTGGAGGATCGCT 4039
QY 4021 TGAGCCCTGGGAGGTGAAGACTGCGAGTGTGATTTGTACCAAGCCCTTAGCTGGG 4080
Db 4040 TGAGCCCTGGGAGGTGAAGACTGCGAGTGTGATTTGTACCAAGCCCTTAGCTGGG 4099
QY 4081 GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAATTTGCAAAAGTGAATTAAGAGT 4140
Db 4100 GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAATTTGCAAAAGTGAATTAAGAGT 4159
QY 4141 GCGTATATGCTAGGCGGAGTGGCTCATGGCTGTAAATCCGACACTTTGGGAAGCCGAG 4200
Db 4160 GCGTATATGCTAGGCGGAGTGGCTCATGGCTGTAAATCCGACACTTTGGGAAGCCGAG 4219
QY 4201 GCGGGGGGCTACCTTAAGTCAAGAGTGTGAGACAGCTGGCCAAACATGGAGAAAGCC 4260
Db 4220 GCGGGGGGCTACCTTAAGTCAAGAGTGTGAGACAGCTGGCCAAACATGGAGAAAGCC 4279
QY 4261 ATCTCTTAATAAATAAATAAATTAAGCGGGGCTGGGGGCAATGGTGAGCAATGGCTGTAA 4320
Db 4280 ATCTCTTAATAAATAAATAAATTAAGCGGGGCTGGGGGCAATGGTGAGCAATGGCTGTAA 4339
QY 4321 TCCGAGCTACTCAGAGAGCTGAGGCGAGAGATCACTTGAACCCAGAGGCGGCGTTCG 4380
Db 4340 TCCGAGCTACTCAGAGAGCTGAGGCGAGAGATCACTTGAACCCAGAGGCGGCGTTCG 4399
QY 4381 AGTGAGCCGAGATGTCCTCATTTGCATTCACCTCCAGCTGGGCAACAGAGCCAAA 4440
Db 4400 AGTGAGCCGAGATGTCCTCATTTGCATTCACCTCCAGCTGGGCAACAGAGCCAAA 4459
QY 4441 CTCTGTCTTAATAAATAAATAAATTAAGCGGGGCTGGGGGCAATGGTGAGCAATGGCTGTAA 4500
Db 4460 CTCTGTCTTAATAAATAAATAAATTAAGCGGGGCTGGGGGCAATGGTGAGCAATGGCTGTAA 4519
QY 4501 TTGCGAGGCAACATGTTTAAGAAATGTGAGGCTCCTGGCTTCATATGCTCTTTAAAAAC 4560
Db 4520 TTGCGAGGCAACATGTTTAAGAAATGTGAGGCTCCTGGCTTCATATGCTCTTTAAAAAC 4579
QY 4561 CACCCTTAAGGCGCAGGTGCGATGCTCATGCTTAATCCAGCACTTTGGGAGCGGAG 4620
Db 4580 CACCCTTAAGGCGCAGGTGCGATGCTCATGCTTAATCCAGCACTTTGGGAGCGGAG 4639
QY 4621 GCGGGTGGATCACCCTGAGGCTGAGAGTTCGAGACAGGCTGACCAACCAATGGTGAAT 4680
Db 4640 GCGGGTGGATCACCCTGAGGCTGAGAGTTCGAGACAGGCTGACCAACCAATGGTGAAT 4699
QY 4681 CCCACCTCTACTAATAAATAAATAAATTAAGATGAGCATGGTGTGATGCCCTTAATCCAC 4740
Db 4700 CCCACCTCTACTAATAAATAAATAAATTAAGATGAGCATGGTGTGATGCCCTTAATCCAC 4759
QY 4741 CTACTTGGGAGGCTGAGGCGAATAATCACTAGAACAGGAGCGGAGGTTGATGTAG 4800
Db 4760 CTACTTGGGAGGCTGAGGCGAATAATCACTAGAACAGGAGCGGAGGTTGATGTAG 4819
QY 4801 CCGGATGCTGTCCTGACTCCAGGCTGAGCAATGAGCGAAATCCACTCAAAAAAAC 4860
Db 4820 CCGGATGCTGTCCTGACTCCAGGCTGAGCAATGAGCGAAATCCACTCAAAAAAAC 4879
QY 4861 AACAAACAAAAACCCACTCTACTCCAGGAGGCTGGGTACAGAGCTGGGCCACATCAGT 4920
Db 4880 AACAAACAAAAACCCACTCTACTCCAGGAGGCTGGGTACAGAGCTGGGCCACATCAGT 4939
QY 4921 GCAAGGTGCTGAGGCGCAGAGACTTAAGGCGAGCTGCGAGAGCCGGGAGCAATTAACAGT 4980
Db 4940 GCAAGGTGCTGAGGCGCAGAGACTTAAGGCGAGCTGCGAGAGCCGGGAGCAATTAACAGT 4999
QY 4981 TGTAGATCACTGTGTGATCAGAGCTCCCTGCATTTGTGACACACAGGAGGAGCCCA 5040
Db 5000 TGTAGATCACTGTGTGATCAGAGCTCCCTGCATTTGTGACACACAGGAGGAGCCCA 5059
QY 5041 AGCACAGAGATGGCCCATCAGTCAACACATCACTTCTCATCCAGAGATGTCTGTTT 5100
Db 5060 AGCACAGAGATGGCCCATCAGTCAACACATCACTTCTCATCCAGAGATGTCTGTTT 5119

QY 5101 CTGGGACGCTGGGGTAATTAAGACAGAGGTGACACTTTGGGTGTGTCAGTCAGAC 5160
Db 5120 CTGGGACGCTGGGGTAATTAAGACAGAGGTGACACTTTGGGTGTGTCAGTCAGAC 5179
QY 5161 TGCCCGCAGGAGGCTTGTGCTGTAGAAAAAGCTTCAAGGCTTAGGCGGCGAGGTGGC 5220
Db 5180 TGCCCGCAGGAGGCTTGTGCTGTAGAAAAAGCTTCAAGGCTTAGGCGGCGAGGTGGC 5239
QY 5221 TCAGGCTGTAAATCCACACTTTGGAGGCGGAGGCGGCTGATCAGAGTCAAGAGA 5280
Db 5240 TCAGGCTGTAAATCCACACTTTGGAGGCGGAGGCGGCTGATCAGAGTCAAGAGA 5299
QY 5281 TCGTGACCATCTGCTGAACAGGTGAAGAAACCCGCTCTACTAAAAATACAAAAATTTGG 5340
Db 5300 TCGTGACCATCTGCTGAACAGGTGAAGAAACCCGCTCTACTAAAAATACAAAAATTTGG 5359
QY 5341 CCGGACATGTTGGCGGCGACCTGTAGTTCAGACTCTCGGAGGCTGAGGAGAGAAATG 5400
Db 5360 CCGGACATGTTGGCGGCGACCTGTAGTTCAGACTCTCGGAGGCTGAGGAGAGAAATG 5419
QY 5401 GCGTGAACCCGAGAGGCGAGAGTTCAGTGAAGCCGAGATCGCGCACTGCACTCCAGCT 5460
Db 5420 GCGTGAACCCGAGAGGCGAGAGTTCAGTGAAGCCGAGATCGCGCACTGCACTCCAGCT 5479
QY 5461 GGGGACAGAGCAAGACTCCATCTGGAAGAAAAAGAAAAAGTTCAAGTGTGAGCCAGA 5520
Db 5480 GGGGACAGAGCAAGACTCCATCTGGAAGAAAAAGAAAAAGTTCAAGTGTGAGCCAGA 5539
QY 5521 GGGCCAGGCTGAATTTGTCACTTACCATGACCTTGGGCAAGGACCTCTCCCTGGC 5580
Db 5540 GGGCCAGGCTGAATTTGTCACTTACCATGACCTTGGGCAAGGACCTCTCCCTGGC 5599
QY 5581 CCACTTCAAGGAGTGAATGACATCACTCAAGGTCCTTCAGACATTAAGCTGATGTTG 5640
Db 5600 CCACTTCAAGGAGTGAATGACATCACTCAAGGTCCTTCAGACATTAAGCTGATGTTG 5659
QY 5641 TAAATGGAAGATGGGAGCATTTCCCTCTCTACACCCAGCCGCTGTCACTTCAAGT 5700
Db 5660 TAAATGGAAGATGGGAGCATTTCCCTCTCTACACCCAGCCGCTGTCACTTCAAGT 5719
QY 5701 GAATGACAGGAGAAATCACTGTCCTCAATCCCGAGTTCCAAAGGCTTGGGAGACCTTAC 5760
Db 5720 GAATGACAGGAGAAATCACTGTCCTCAATCCCGAGTTCCAAAGGCTTGGGAGACCTTAC 5779
QY 5761 TGTCAAGGTGCTGACAGAGAGTGAAGTCAAGTCAAGTCAAGTCCCTGGAAGGCTTGG 5820
Db 5780 TGTCAAGGTGCTGACAGAGAGTGAAGTCAAGTCAAGTCAAGTCCCTGGAAGGCTTGG 5839
QY 5821 CCTCATTTGGGACAGACATCCGCTTCTCTGCGCTCAACCGGATTTAGGGGCTTACG 5880
Db 5840 CCTCATTTGGGACAGACATCCGCTTCTCTGCGCTCAACCGGATTTAGGGGCTTACG 5899
QY 5881 GCAATGAGTCAATGGGAGGCGGAGGCTTCTGGGAGATTTCCAGCTTAATCAACTTGGGA 5940
Db 5900 GCAATGAGTCAATGGGAGGCGGAGGCTTCTGGGAGATTTCCAGCTTAATCAACTTGGGA 5959
QY 5941 CAGGACAGCTGGAACCTTGTGATGGTGCCTATCCAAATGTGGGGTGGGACAGCCAA 6000
Db 5960 CAGGACAGCTGGAACCTTGTGATGGTGCCTATCCAAATGTGGGGTGGGACAGCCAA 6019
QY 6001 GACCAATGTCTTATATCAAGTAGGGCTCAAGAGTCTCCAGACAGGACCTCCGG 6060
Db 6020 GACCAATGTCTTATATCAAGTAGGGCTCAAGAGTCTCCAGACAGGACCTCCGG 6079
QY 6061 AGAGTTTGGGGTGAAGTGAAGCAACAGGCTCTTTTCTTCTTGAATTTGGG 6120
Db 6080 AGAGTTTGGGGTGAAGTGAAGCAACAGGCTCTTTTCTTCTTGAATTTGGG 6139
QY 6121 GCGTTGGGGAAGGCTTGAAGATCCAAAGAGAGGGAAGAGCACTCCCCACAG 6180
Db 6140 GCGTTGGGGAAGGCTTGAAGATCCAAAGAGAGGGAAGAGCACTCCCCACAG 6199

QY		6181	TCTGCGCAGACGAGAGAGAGACCCTGCACCTTCCTCCACAGGCTT	6235
Dd		6200	TCTGCCAGACGAGAGAGAGACCCTGCACCTTCCTCCACAGGCTT	6254
RESULT 5				
AC090844/C				
LOCUS	AC090844	227857 bp	DNA	linear PRI 29-AUG-2002
DEFINITION	Homo sapiens chromosome 17, clone RP11-387H17, complete sequence.			
ACCESSION	AC090844			
VERSION	AC090844.7	GI:22539183		
KEYWORDS	HTG.			
SOURCE				
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 227857)			
AUTHORS	Birren,B., Nusbaum,C. and Lander,E.			
TITLE	Homo sapiens chromosome 17, clone RP11-387H17			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 227857)			
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,I., Bouckgalter,B., Brown,A., Camaraeta,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgeraid,M., Gage,D., Galagan,J., Gardayna,S., Glnde,S., Goylete,M., Graham,L., Grand-Pierre,N., Hagos,B., Harford,A., Horton,L., Hulme,W., Illev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K., Lamazars,R., Lander,T., Lenoczky,J., Levine,R., Liu,G., Maclean,C., McDonald,P., Marguez,N., Matthews,C., McCarthy,M., McKernan,K., McKeeeters,R., Meldrim,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Ribback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnuppach,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,T., Travers,M., Travis,N., Triggilo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (11-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 227857)			
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,I., Bouckgalter,B., Camaraeta,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgeraid,M., Gage,D., Galagan,J., Gardayna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Illev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., McDonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schnuppach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J., Teefaye,S., Theodore,T., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	4 (bases 1 to 227857)			
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,I., Bouckgalter,B., Camaraeta,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgeraid,M., Gage,D., Galagan,J., Gardayna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Illev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., McDonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schnuppach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J., Teefaye,S., Theodore,T., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			

TITLE	Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome Research 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL	On Aug 29, 2002 this sequence version replaced g1:22213295.
COMMENT	All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L12018 Center clone name: 387_H_17 -----
FEATURES	Location/Qualifiers 1..227857 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="RP11-387H17" /clone_id="RPC1-11 Human Male BAC" complement(4..40) /rpt_family="LMB8" 41..348 /rpt_family="Alusx" complement(349..456) /rpt_family="LMB8" complement(514..824) /rpt_family="Alusx" 3751..3913 /rpt_family="MIR" complement(3961..4285) /rpt_family="Alusx" complement(4288..4570) /rpt_family="AluJo" 5865..5908 /rpt_family="(TG)n" 6880..7112 /rpt_family="L2" complement(9578..9582) /note"<30 qual SNGI region" complement(9650..9654) /note"<30 qual SNGI region" complement(10553..10686) /rpt_family="MIR3" complement(10698..10994) /rpt_family="AluY" complement(11328..11629) /rpt_family="AluSg" complement(11733..12049) /rpt_family="AluSp" complement(12097..12319) /rpt_family="AluJb" complement(12331..12944) /rpt_family="prr5" complement(12764..12773) -----
SOURCE	

```

repeat_region      /note="<30 qual SNGL region"
12945. .13393
/rpl_family="LTR7"
complement(13154. .13158)
unsure             /note="<30 qual SNGL region"
13293. .13339
/note="single clone coverage"
unsure             13307. .13311
/note="<30 qual single clone coverage"
unsure             13381. .13386
/note="<30 qual SNGL region"
repeat_region      complement(13394. .14319)
/rpl_family="PTR5"
unsure             13402. .13407
/note="<30 qual SNGL region"
repeat_region      complement(14320. .14385)
/rpl_family="Aluub"
repeat_region      complement(14396. .14552)
/rpl_family="MIR3"
repeat_region      14855. .15000
/rpl_family="L1ME"
repeat_region      complement(15072. .15371)
/rpl_family="Alusx"
repeat_region      15372. .15581
/rpl_family="MER58B"
repeat_region      complement(15637. .15814)
/rpl_family="Aluub"
repeat_region      complement(15815. .16120)
/rpl_family="Alusx"
repeat_region      complement(16121. .16257)
/rpl_family="Aluub"
repeat_region      16447. .16478
/rpl_family="L2"
repeat_region      16580. .16863
/rpl_family="AluO"
repeat_region      16870. .16896
/rpl_family="(TAAA)n"
repeat_region      16899. .17217
/rpl_family="MSTA"
repeat_region      complement(17244. .17412)
/rpl_family="Aluub"
repeat_region      complement(17413. .17709)
/rpl_family="AlusC"
repeat_region      complement(17711. .18000)
/rpl_family="Aluy"
repeat_region      complement(18001. .18132)
/rpl_family="Aluub"
repeat_region      18219. .18259
/rpl_family="L1MC4"
repeat_region      complement(18360. .18386)
/rpl_family="FLAM_C"
repeat_region      18387. .18526
/rpl_family="L1MC4"
repeat_region      18541. .18845
/rpl_family="Alusx"
repeat_region      18897. .19197
/rpl_family="Alusx"
repeat_region      19637. .20125
/rpl_family="MLT2A2"
repeat_region      complement(20626. .20736)
/rpl_family="Tigger1"

```

Query Match 97.1% Score 6051.2; DB 9; Length 227857;
 Best Local Similarity 99.4% Pred. No. 0;
 Matches 6201; Conservative 0; Mismatches 23; Indels 16; Gaps 12;

```

QY 1 GATCATTGAGAGAGTATGATCAAGACGAGCTGGGAGCATAGGAGAGTGTCTACG 60
DB 16197 GATCATTGAGAGAGTATGATCAAGACGAGCTGGGAGCATAGGAGAGTGTCTACG 16138
QY 61 AAAAAATCAAAATTAATGCGCGGAGCATGCTGCTACGCTGTATATCCCTGAACCTTTGGG 120
DB 16137 AAAAAATCAAAATTAATGCGCGGAGCATGCTGCTACGCTGTATATCCCTGAACCTTTGGG 16078

```

```

QY 121 ACATCAGGCAAGTATGATCATTGGAGTCAGAGTTCGAGACTAGCTGGCCACATGCT 180
DB 16077 ACATCAGGCAAGTATGATCATTGGAGTCAGAGTTCGAGACTAGCTGGCCACATGCT 16018
QY 181 GAAACCTATCTCCATAAAAAATCAAAATAATACAGGCAATGCTGGCAGGCACTGTGA 240
DB 16017 GAAACCTATCTCTACT-AAAAATCAAAATAATACAGGCAATGCTGGCAGGCACTGTGA 15959
QY 241 ATCCGGCTACTCAGAGGCTGAGGAGAGATACCTTGAACCCAGAGGCGGAGGTTG 300
DB 15958 ATCCGGCTACTCAGAGGCTGAGGAGAGATACCTTGAACCCAGAGGCGGAGGTTG 15899
QY 301 CAGTACGCTGATACACCCACTGACCTCCAGCCGGGGGAGAGCAAGATCATATC 360
DB 15898 CAGTACGCTGATACACCCACTGACCTCCAGCCGGGGGAGAGCAAGATCATATC 15839
QY 361 AAAAAAATTAATAAAAAATTAAGCCAGCATGAGTACACCTGAGTCTGAG 420
DB 15838 AAAAAAATTAATAAAAAATTAAGCCAGCATGAGTACACCTGAGTCTGAG 15779
QY 421 CTACTCAGAGGCTGAGTGGGAGATCATTTGAACCTGGGCACTCAAGGCTACAGTGA 480
DB 15778 CTACTCAGAGGCTGAGTGGGAGATCATTTGAACCTGGGCACTCAAGGCTACAGTGA 15720
QY 481 GCCAAGATCATGCCACTACCTCCAGCCTGGGCAACAGAGAGAGACCTGTCTATAAAA 540
DB 15719 GCCAAGATCATGCCACTACCTCCAGCCTGGGCAACAGAGAGACCTGTCTATAAAA 15660
QY 541 AATATATTAATTAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 15659 AATATATTAATTAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15603
QY 601 ATGTATATAGTTTGCAAACTCAAGATCAGATGATGATTTTAAAGCTTGGCCCT 660
DB 15602 ATGTATATAGTTTGCAAACTCAAGATCAGATGATGATTTTAAAGCTTGGCCCT 15543
QY 661 ATGCTCTGTCACAAATCAGCTGCGGCTGCTTTTACACAAAGAGAGATTAACAT 720
DB 15542 ATGCTCTGTCACAAATCAGCTGCGGCTGCTTTTACACAAAGAGAGATTAACAT 15483
QY 721 ACATCAGATGAATTTTATAGACATCGAGATTTGAATTCATATGATTTTACATTTAT 780
DB 15482 ACATCAGATGAATTTTATAGACATCGAGATTTGAATTCATATGATTTTACATTTAT 15423
QY 781 AAAAAATCTTTTAAAAATTTTCCCTACCATTTTAAAAAGTGAAGCGGCCA---G 837
DB 15422 AAAAAATCTTTTAAAAATTTTCCCTACCATTTTAAAAAGTGAAGCGGCCA---G 15363
QY 838 CGCGCATCGTCACGCGCTGTAATTCACAGCTTTGGGAGGCTGAGTGGGAGATCAGT 897
DB 15362 CGCGCATCGTCACGCGCTGTAATTCACAGCTTTGGGAGGCTGAGTGGGAGATCAGT 15303
QY 898 GAGATCAACAGTTCGAGACAGCTGGCCCAACATAGCAAAACCCATTTCTATAAAT 957
DB 15302 GAGATCAACAGTTCGAGACAGCTGGCCCAACATAGCAAAACCCATTTCTATAAAT 15243
QY 958 AAAAAATTAAGTGGGCACTAGTGTGCAACACTGTGATCCAGCTACTTGGGAGGCTGAG 1017
DB 15242 AAAAAATTAAGTGGGCACTAGTGTGCAACACTGTGATCCAGCTACTTGGGAGGCTGAG 15183
QY 1018 GCAGAGAAATGCTTGAACCTGGGAGAGGAGGAGTTCAGTGAACATCATATGCTACATG 1077
DB 15182 GCAGAGAAATGCTTGAACCTGGGAGAGGAGGAGTTCAGTGAACATCATATGCTACATG 15123
QY 1078 CACTCAGACCTGGGAGAGAGTGAAGTTCGCTCAAGC-AAAAAATAAATGTAAGAAG 1136
DB 15122 CACTCAGACCTGGGAGAGAGTGAAGTTCGCTCAAGC-AAAAAATAAATGTAAGAAG 15063
QY 1137 CCATTCTTAATTCAGTGTACATCAGTGTACATCTCAGGCTGCTGCTGCTGAG 1196
DB 15062 CCATTCTTAATTCAGTGTACATCAGTGTACATCTCAGGCTGCTGCTGCTGAG 15003

```

QY 1197 GCATACCTGAGAGTAGAGTTGCTTGCTACAGAGCAATACACTTTCCATTACTAGA 1256
|||||
Db 15002 GCATACCTGAGAGTAGAGTTGCTTGCTTGCTACAGAGCAATACACTTTCCATTACTAGA 14943
QY 1257 CACTACCAAGTTCCATCCAAAGAGGTTTTTTTTTACAAATCTACACTCCCCAGAAC 1316
|||||
Db 14942 CACTACCAAGTTCCATCCAAAGAGGTTTTTTTTTACAAATCTACACTCCCCAGAAC 14883
QY 1317 AATGAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAAC 1376
14882 AATGAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAAC 14823
QY 1377 AGAATGAGAGGAGGAGAGCTGCCAGCCCTTCTAACCATGAAATAATCTGGTAGACC 1436
14822 AGAATGAGAGGAGGAGAGCTGCCAGCCCTTCTAACCATGAAATAATCTGGTAGACC 14763
QY 1437 TTCTGATGCTGGAAAGATGAAATAAGGGGGTCTGAGAGCTGCCCTGTAGATCAC 1496
14762 TTCTGATGCTGGAAAGATGAAATAAGGGGGTCTGAGAGCTGCCCTGTAGATCAC 14703
QY 1497 TGTGACTTGTAGACCTTCAGCTCAGTCTCAGCCCATGTGTCAATGGCCAGTAAATGAG 1556
14702 TGTGACTTGTAGACCTTCAGCTCAGTCTCAGCCCATGTGTCAATGGCCAGTAAATGAG 14644
QY 1557 CCTCAGCTCTGTTGGTCTTATTTCT-CCCCATGAGGGGGTGAAGTGTGATGAGCC 1615
14643 CCTCAGCTCTGTTGGTCTTATTTCT-CCCCATGAGGGGGTGAAGTGTGATGAGCC 14584
QY 1616 GTTATTCAGAGTGTACAGCTTTCTTGACAGAAAGTAGTGTCAAGAAACAGCAGGGGCT 1675
14583 GTTATTCAGAGTGTACAGCTTTCTTGACAGAAAGTAGTGTCAAGAAACAGCAGGGGCT 14524
QY 1676 TGGCAAGATGATTAACCTGCAAAATCCTACCTGGCTCAGCCACAGCTAGTTGTGATCT 1735
14523 TGGCAAGATGATTAACCTGCAAAATCCTACCTGGCTCAGCCACAGCTAGTTGTGATCT 14464
QY 1736 TGAACAAGTTTTTTCACCTCTCTGAGGGCATCCCTGGGCTACAACAACAGTGGTGA 1795
14463 TGAACAAGTTTTTTCACCTCTCTGAGGGCATCCCTGGGCTACAACAACAGTGGTGA 14404
QY 1796 CAGGATGAATGACG- AAGTCCCTTACACCTGTAAATCCAGCACTTTGGAGGCCAAGGC 1854
14403 CAGGATGAATGACGAAAGTGGCTTACACTGTAAATCCAGCACTTTGGAGGCCAAGGC 14344
QY 1855 GGGTGAATGAGCTTGAGACCTGAGAGGTGACAGCATGGCCGACATGTCACAGCCCTGTTTC 1914
14343 GGGTGAATGAGCTTGAGACCTGAGAGGTGACAGCATGGCCGACATGTCACAGCCCTGTTTC 14284
QY 1915 GCTCTGGGCGCTCTCTGCTGGGCTCCACATTGGGTGACATTGAGAGCCCTTACG 1974
14283 GCTCTGGGCGCTCTCTGCTGGGCTCCACATTGGGTGACATTGAGAGCCCTTACG 14224
QY 1975 CCACGCTGCACTGTGGGAGCCCTTTCTGGGCTGGCCAAAGGCCAGCCGGCTCCCTCA 2034
14223 CCACGCTGCACTGTGGGAGCCCTTTCTGGGCTGGCCAAAGGCCAGCCGGCTCCCTCA 14164
QY 2035 GCTTGAAGAGAGTGTGAGAGGAGGCTTAACAGAGAACCGGGGCTGCCACAGGGCTT 2094
14163 GCTTGAAGAGAGTGTGAGAGGAGGCTTAACAGAGAACCGGGGCTGCCAGGGCTT 14104
QY 2095 GCGGGCAGCTGAGATTCCGGGTGGGGCTGGGCTGGGGGGGCCCGACATCGAGAGCAGC 2154
14103 GCGGGCAGCTGAGATTCCGGGTGGGGCTGGGCTGGGGGGGCCCGACATCGAGAGCAGC 14044
QY 2155 GGGCAGCCCTGCGAGGCCCCGGGCAATGAGAGGCTTAAAGCCCGGGCAGAGGGCTCGGA 2214
14043 GGGCAGCCCTGCGAGGCCCCGGGCAATGAGAGGCTTAAAGCCCGGGCAGAGGGCTCGGA 13984
QY 2215 GGGTGAATGAGTGTGCCCCAGACAGTGCACGCCCGGGCTGTGCTCGATTTCTCA 2274
13983 GGGTGAATGAGTGTGCCCCAGACAGTGCACGCCCGGGCTGTGCTCGATTTCTCA 13924
QY 2275 CTGGGCTTAAAGAGCTTCCCGGGGGCAGAGGCTCGGGAGCTCAGAGCCCGCATGCTGA 2334

Db 13923 CTGGGCTTAAAGAGCTTCCCGGGGGCAGGGCTCGGAGCTCAGAGCCCGCATGCTGA 13864
QY 2335 GCCCTCCCTCAATGAGGCTCTCTGCGGCCCGAGACCTCCCGAGAGACACCCCGTGT 2394
13863 GCCCTCCCTCAATGAGGCTCTCTGCGGCCCGAGACCTCCCGAGAGACACCCCGTGT 13804
QY 2395 CCACAGCGCCAGTCCCATGACACAGCAAGGCTGAGAAATGCGGGGCGACGGCACCGG 2454
13803 CCACAGCGCCAGTCCCATGACACAGCAAGGCTGAGAAATGCGGGGCGACGGCACCGG 13744
QY 2455 GACTGGAGGCACTACCCCTGACAGCCCTGGTGGGAATCCTGGGTGAAGCAGCTGG 2514
13743 GACTGGAGGCACTACCCCTGACAGCCCTGGTGGGAATCCTGGGTGAAGCAGCTGG 13684
QY 2515 GCTCTGAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGAGGATGTAATATAC 2574
13683 GCTCTGAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGAGGATGTAATATAC 13624
QY 2575 ACCAATCAGACCCCTGTGTCTAGCTCAGGCTGTGTGAATCCACACTCTGTAT 2634
13623 ACCAATCAGACCCCTGTGTCTAGCTCAGGTTCTGTGAATCCACACTCTGTAT 13564
QY 2635 CTAAGCTACTGTGAATGGGCGCTTGAGAACCTTATGTCTAGCTCAGAGGATGTAATATCA 2694
13563 CTAAGCTACTGTGAATGGGCGCTTGAGAACCTTATGTCTAGCTCAGAGGATGTAATATCA 13504
QY 2695 CCAATCGGCACTGTATCTAGCTCAAGGTTGTAAACACACCAATCAGACCCGTGTTC 2754
13503 CCAATCGGCACTGTATCTAGCTCAAGGTTGTAAACACACCAATCAGACCCGTGTTC 13444
QY 2755 TAGCTAGGGTATGTGAATGACCAATCGACAGTGTATCTGTGCTACTTCAATGGCAT 2814
13443 TAGCTAGGGTATGTGAATGACCAATCGACAGTGTGTATCTGTGCTACTTCAATGGCAT 13384
QY 2815 CCGTGTGAAGAGACACCAACAGAGCTTGTGTGAGCAATAAAGCTTCATACCTGGGT 2874
13383 CCGTGTGAAGAGACACCAACAGAGCTTGTGTGAGCAATAAAGCTTTCATACCTGGGT 13324
QY 2875 GCAAGTGGGCTGAGTCCGAAAGAGAGTCAAGCAAGGAGATAAAGGTGGGCGCTTTTA 2934
13323 GCAAGTGGGCTGAGTCCGAAAGAGAGTCAAGCAAGGAGATAAAGGTGGGCGCTTTTA 13264
QY 2935 TAGGATTTGGGTAGTAAAGAAATTAACAGTCAAAAGGGGTTGTCTCTGCGGGCAG 2994
13263 TAGGATTTGGGTAGTAAAGAAATTAACAGTCAAAAGGGGTTGTCTCTGCGGGCAG 13204
QY 2995 GAGTGGGGGCTGCAAGGTCTCAGTGGGGGCTTTTGTAGCCAGATGAGCCAGAAA 3054
13203 GAGTGGGGGCTGCAAGGTCTCAGTGGGGGCTTTTGTAGCCAGATGAGCCAGAAA 13144
QY 3055 AGGACTTTCACAAAGTAAATGTCAATTAAGCAAGAGACCCGCATTTACACTCTTTT 3114
13143 AGGACTTTCACAAAGTAAATGTCAATTAAGCAAGAGACCCGCATTTACACTCTTTT 13084
QY 3115 GTGGTGAATGATCATCAATTAAGTGGGAGGAGCAATTCATCTCTTTGTATTTTC 3174
13083 GTGGTGAATGATCATCAATTAAGTGGGAGGAGCAATTCATCTCTTTGTATTTTC 13024
QY 3175 AGTTACTTTCAGGCATCTGGGCGTAAATGTGCAAGTTTACAGGGGATGCAATGGCTGGCT 3234
13023 AGTTACTTTCAGGCATCTGGGCGTAAATGTGCAAGTTTACAGGGGATGCAATGGCTGGCT 12964
QY 3235 TGGGCTCAGAGGCTTGAACAGCTACTGTGTGGGGCTTTGGAGATGTTTGTGACACT 3294
12963 TGGGCTCAGAGGCTTGAACAGCTACTGTGTGGGGCTTTGGAGATGTTTGTGACACT 12904
QY 3295 CTGTATCTAGTTAATCTAGTGGGAGGTGAGAACTTTGTGTCTAGCTCAGGGAATGTA 3354
12903 CTGTATCTAGTTAATCTAGTGGGAGGTGAGAACTTTGTGTCTAGCTCAGGGAATGTA 12844
QY 3355 AAGCAACCAATCAGCCCGCTGTCAAAAACAGACACTGCGCTCAACCAATCAGAGATGT 3414
|||||

Db 12843 AACGACCAATACAGGCGCCCTGTCAAAACAGCACTCGGCTCTTACCAATACAGCATGT 12784
Oy 3415 GGGTGGGCGCAGATTAAGAAATAAAGAGGCTGCCGAGCGCAGCATGGCAAGCGGAC 3474
Db 12783 GGGTGGGCGCAGATTAAGAAATAAAGAGGCTGCCGAGCGCAGCATGGCAAGCGGAC 12724
Oy 3475 AGTCCCTATCCACAATATGGCAGCTTTGTCTTTGCTGTTTGGCATTAATCTTGTCAC 3534
Db 12723 AGGTCCCATCCACAATATGGCAGCTTTGTCTTTGCTGTTTGGCATTAATCTTGTCAC 12664
Oy 3535 TGTCTGCTTTTGGGTCCACACTGCTTTTATAGCTGTAACTACACACGAGGCTTGC 3594
Db 12663 TGTCTGCTTTTGGGTCCACACTGCTTTTATAGCTGTAACTACACACGAGGCTTGC 12604
Oy 3595 AGCTTCACTCCGAAAGCCACTAAGCCACGAGCCCGGAGGAGAAATGAACACTCCGGC 3654
Db 12603 AGCTTCACTCCGAAAGCCACTAAGCCACGAGCCCGGAGGAGAAATGAACACTCCGGC 12544
Oy 3655 CGCGTGCCTTAAGAGCTATTAACACTCACCGGAGGTCTGCAGCTTCACTCTCAGGCA 3714
Db 12543 CGCGTGCCTTAAGAGCTATTAACACTCACCGGAGGTCTGCAGCTTCACTCTCAGGCA 12484
Oy 3715 GCGAGACCAAGAACCCACAGAGAGAAACTCGGAACACATCTGAACATCAGAGAA 3774
Db 12483 GCGAGACCAAGAACCCACAGAGAGAAACTCGGAACACATCTGAACATCAGAGAA 12424
Oy 3775 CAACTCCAGATGCAACCCCTTAAGAGCTGTAACTACACTACGAGGCTCCGGCTTCC 3834
Db 12423 CAACTCCAGATGCAACCCCTTAAGAGCTGTAACTACACTACGAGGCTCCGGCTTCC 12364
Oy 3835 TTCTTGAAGTACGTAGACCAAGCACTCACCAGTTTGGACACAGCCAGAGTTTACG 3894
Db 12363 TTCTTGAAGTACGTAGACCAAGCACTCACCAGTTTGGACACAGCCAGAGTTTACG 12304
Oy 3895 ATCAGCCTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAATAATTACAAAT 3954
Db 12303 ATCAGCCTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAATAATTACAAAT 12245
Oy 3955 TGGGGGAGCATGTGTGTCGCGCTGTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGG 4014
Db 12244 TGGGGGAGCATGTGTGTCGCGCTGTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGG 12185
Oy 4015 ATTCCTTGAAGCTGGGAGGAGCAAGCACTCAGTGAAGCTGTGATTGTACACAGCCCTACG 4074
Db 12184 ATTCCTTGAAGCTGGGAGGAGTAAAGCTCAGTGAAGCTGTGATTGTACACAGCCCTACG 12125
Oy 4075 GCTGGGGGACAGACTGAGACCCCTGTTCCCTCCGCAAAAAAATAATTGACAAAAAGTGAATA 4134
Db 12124 GCTGGGGGACAGACTGAGACCCCTGTTCCCTCCGCAAAAAAATAATTGACAAAAAGTGAATA 12065
Oy 4135 AGAGGTGCTATATATGGTATAGGCGCAGTGGCTCATGCTGTAAATCCACAGCACTTTGGGAA 4194
Db 12064 AGAGGTGCTATATATGGTATAGGCGCAGTGGCTCATGCTGTAAATCCACAGCACTTTGGGAA 12005
Oy 4195 GCCAGAGGCGGCGGTCACTAAGGTGAGAGTGTGAGACAGCCTGGCCAAACATGAGGA 4254
Db 12004 GCCAGAGGCGGCGGTCACTAAGGTGAGAGTGTGAGACAGCCTGGCCAAACATGAGGA 11945
Oy 4255 AAGCCATCTCTTAAAAAATAAAATTAACCGGCTGTGGGGGAGTGTGAGCATGC 4314
Db 11944 AAGCCATCTCTTAAAAAATAAAATTAACCGGCTGTGGGGGAGTGTGAGCATGC 11885
Oy 4315 CTGTAAATCCAGCTATCTCAGAGAGCTGAGGCAAGAAATCACTTGAACCCAGAGGCGGC 4374
Db 11884 CTGTAAATCCAGCTATCTCAGAGAGCTGAGGCAAGAAATCACTTGAACCCAGAGGCGGC 11825
Oy 4375 GGTTCAGTACGAGAGCTGTGCATGTGACCTCAGCCATCCAGCCTTGCGCAACAAGA 4434
Db 11824 GGTTCAGTACGAGAGCTGTGCATGTGACCTCAGCCATCCAGCCTTGCGCAACAAGA 11765
Oy 4435 GCCAAACTCTGTCTTAAAAAATAAAAGTGGCTGACATATAGAGGTGTGCAATG 4494
Db 11764 GCCAAACTCTGTCTTAAAAAATAAAAGTGGCTGACATATAGAGGTGTGCAATG 11705

Oy 4495 CAATAGTTCAGGACACATGTTTAAGAAATGAGAGCTCCGCTTCCATGTCCTGTTA 4554
Db 11704 CAATAGTTCAGGACACATGTTTAAGAAATGAGAGCTCCGCTTCCATGTCCTGTTA 11645
Oy 4555 AAAACCAACCTCAAGGCGCAGGTGACAGTGGCTCATATTAATCCAGCACTTTGGGAG 4614
Db 11644 AAAACCAACCTCAAGGCGCAGGTGACAGTGGCTCATATTAATCCAGCACTTTGGGAG 11585
Oy 4615 GCCGAGGCGGTGATCACCTGAGGTCAAGAGTGTGAGACAGCCTGACCAACATGG 4674
Db 11584 GCCGAGGCGGTGATCACCTGAGGTCAAGAGTGTGAGACAGCCTGACCAACATGG 11525
Oy 4675 TGAATCCACCTCACTAAAAATCAAAATTAAGATGAGAGCTGGTGGCATCCCTGTA 4734
Db 11524 TGAATCCACCTCACTAAAAATCAAAATTAAGATGAGAGCTGGTGGCATCCCTGTA 11465
Oy 4735 TCCCACTTACTTTGGAGGCTGAGGCGAAGAAATCACTAGAACCCAGGAGCGGAGTTGT 4794
Db 11464 TCCCACTTACTTTGGAGGCTGAGGCGAAGAAATCACTAGAACCCAGGAGCGGAGTTGT 11405
Oy 4795 AGTGAGCGGAGATCGTGCATTTGCACTCCAGCCTGAGCAATGAGCGAAATCCATCTCA 4854
Db 11404 AGTGAGCGGAGATCGTGCATTTGCACTCCAGCCTGAGCAATGAGCGAAATCCATCTCA 11345
Oy 4855 AAAAACCAACCAAAACCCACTCTCTACTCCAGGAGCTGGGTACAGAGCTGGGCCAC 4914
Db 11344 AAAAACCAACCAAAACCCACTCTCTACTCCAGGAGCTGGGTACAGAGCTGGGCCAC 11285
Oy 4915 ATCAGTGGCAAGGTGTGAGGCCACAGAGCTTAAGGGGAGCTGAGAGCCGCGGACAGATA 4974
Db 11284 ATCAGTGGCAAGGTGTGAGGCCACAGAGCTTAAGGGGAGCTGAGAGCCGCGGACAGATA 11225
Oy 4975 ACAGTGTGAGATCAGTGTGTGAGATCAGAGCTCCCTGCCATTGGTGACCAACGAGGG 5034
Db 11224 ACAGTGTGAGATCAGTGTGTGAGATCAGAGCTCCCTGCCATTGGTGACCAACGAGGG 11165
Oy 5035 CCCCCAAGCAGCAGATGCCCCCATCCAGTACCACTCCACTTCTCATCCAGAGATGT 5094
Db 11164 CCCCCAAGCAGCAGATGCCCCCATCCAGTACCACTCCACTTCTCATCCAGAGATGT 11105
Oy 5095 CTGTTTCTTGGACCTGGGAGTAAATTAAGACAGAGGAGACAGTGGGTGGTGTACG 5154
Db 11104 CTGTTTCTTGGACCTGGGAGTAAATTAAGACAGAGGAGTAAAGTGTGGTGTGTACG 11045
Oy 5155 TCAGACTGCCCGAGGAGGCTTGTGGCTGTAGAAAAAGTTACAGCCTAGGCGGGCAC 5214
Db 11044 TCAGACTGCCCGAGGAGGCTTGTGGCTGTAGAAAAAGTTACAGCCTAGGCGGGCAC 10985
Oy 5215 GGTGGCTACGCTGTATCCAGACACTTTGGAGAGCCGAGGCGGTGATCAGAGTGC 5274
Db 10984 GGTGGCTACGCTGTATCCAGACACTTTGGAGAGCCGAGGCGGTGATCAGAGTGC 10925
Oy 5275 AGGAGATGTGACATCCGCTTAACAGGAGTAAACCCGCTCTACTAAAAATCAAAA 5334
Db 10924 AGGAGATGTGACATCCGCTTAACAGGAGTAAACCCGCTCTACTAAAAATCAAAA 10865
Oy 5335 AATTGGCCGGGAGTGGTGGCGGCACTGTAGTTTCACTCTCAGGAGGCTGAGAGCAG 5394
Db 10864 AATTGGCCGGGAGTGGTGGCGGCACTGTAGTTTCACTCTCAGGAGGCTGAGAGCAG 10805
Oy 5395 AGAATGGCTGAACCCAGAGAGAGATCTGAGAAAGAAAGAAAGAAAGTTCAAGTCTC 10745
Db 10804 AGAATGGCTGAACCCAGAGAGAGATCTGAGAAAGAAAGAAAGAAAGTTCAAGTCTC 10745
Oy 5455 CAGCCTGGGCGCAGAGCAAGATCACTGAGAAAGAAAGAAAGAAAGTTCAAGTCTC 5514
Db 10744 CAGCCTGGGCGCAGAGCAAGATCACTGAGAAAGAAAGAAAGAAAGTTCAAGTCTC 10665
Oy 5515 GCCAGAGGCCAGGCTGTAAATCTGTCACTTACATGACCTTGGGCAAGGACCTTCTTC 5574
Db 10664 GCCAGAGGCCAGGCTGTAAATCTGTCACTTACATGACCTTGGGCAAGGACCTTCTTC 10625

OY	5755	CGTGCCCAAGTTACAGGGGTTGGAAATCGACTCCAAAGTCCCTTCCAGATTAAAGCTGCA	5634	Meneus, L., Mhova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkanth, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Db	10624	CGTGCCCAAGTTACAGGGGTTGGAAATCGACTCCAAAGTCCCTTCCAGATTAAAGCTGCA	10555	
OY	5635	TGCTTCTAAGTAGAAGATGGGGGAGTTTCCCTCTCTCACCCAGCCCGGTGTCCACTT	5694	
Db	10564	TGCTTCTAAGTAGAAGATGGGGGAGTTTCCCTCTCTCACCCAGCCCGGTGTCCACTT	10505	
OY	5695	CAAGTGAATGACCGGGAAGTACAGTGTCCCAATCCCGCACTTCCAAAGCCCTTGGGGA	5754	
Db	10504	CAAGTGAATGACCGGGAAGTACAGTGTCCCAATCCCGCACTTCCAAAGCCCTTGGGGA	10445	
OY	5755	CCCTACTGTCAAGGGTCTGTGACAGAGAGTGAAGTCAAGTGAAGCCCAATCCGCTGGAAG	5814	
Db	10444	CCCTACTGTCAAGGGTCTGTGACAGAGAGTGAAGTCAAGTGAAGCCCAATCCGCTGGAAG	10385	
OY	5815	GTCCTTGCTCATTTCCGAGACAGACATCCGGTTCTCTGTCTTACCGGGATTCTAAGGGC	5874	
Db	10384	GTCCTTGCTCATTTCCGAGACAGACATCCGGTTCTCTGTCTTACCGGGATTCTAAGGGC	10325	
OY	5875	TTTAGCCGAATGAGTCATGGGGGGGGGGGTTTCTGGGGAGTTCCTCAGCTAATCAAC	5934	
Db	10324	TTTAGCCGAATGAGTCATGGGGGGGGGGGTTTCTGGGGAGTTCCTCAGCTAATCAAC	10265	
OY	5935	TTGGACAGAGACAGCTTGAACTTTCATGCTGCTATCAAGTGTGGGTGGGACAGC	5994	
Db	10264	TTGGACAGAGACAGCTTGAACTTTCATGCTGCTATCAAGTGTGGGTGGGACAGC	10205	
OY	5995	AGCCAAAGCCCAATGTCTTATTCAGTACAGGGGCTCAGAGTCTCCACAGACAGCAGC	6054	
Db	10204	AGCCAAAGCCCAATGTCTTATTCAGTACAGGGGCTCAGAGTCTCCACAGACAGCAGC	10145	
OY	6055	CTCCGAGAGATTTGGGGGTAGGAATGGAGACACAGGCTTTTCTCTCTTACAA	6114	
Db	10144	CTCCGAGAGATTTGGGGGTAGGAATGGAGACACAGGCTTTTCTCTCTTACAA	10086	
OY	6115	TTTGGGGGCTTGGGGGACAGGCTTGAGAAATCCCAAGAGAGGGGCAAGAGCACT-CC	6173	
Db	10085	TTTGGGGGCTTGGGGGACAGGCTTGAGAAATCCCAAGAGAGGGGCAAGAGCACTGCC	10026	
OY	6174	CCACAGTCTGCGAGAGAGAGAGAGGAGACCCGACGCTGACACTTCCCAACAGGC	6233	
Db	10025	CCGCAAGTCTGCGAGAGAGAGAGGAGACCCGACGCTGACACTTCCCAACAGGC	9967	
RESULT 6				
AC102799/c		122132 bp	DNA	linear
LOCUS				PRI 01-APR-2003
DEFINITION				Homo sapiens chromosome 17, clone CTD-2248B4, complete sequence.
ACCESSION				AC102799
VERSION				AC102799.9 GI:29423944
KEYWORDS				HTG.
SOURCE				Homo sapiens (human)
ORGANISM				Homo sapiens
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE				1 (bases 1 to 122132)
AUTHORS				Bitren, B., Nussbaum, C. and Lander, E.
TITLE				Homo sapiens chromosome 17, clone CTD-2248B4
JOURNAL				Unpublished
REFERENCE				2 (bases 1 to 122132)
AUTHORS				Bitren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Ferreira, P., Fitzgerald, M., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Girdya, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczký, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
COMMENT				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				

FEATURES		Location/Qualifiers	
Source		1..122132	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="17"	
		/map="17"	
		/clone="CTD-2248E4"	
		/clone_lib="CITD1 Human BAC"	
repeat_region		complement(10..101)	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(95..131)	
repeat_region		/rpt_family="MERSA"	
repeat_region		132..435	
repeat_region		/rpt_family="AluSg"	
repeat_region		complement(436..534)	
repeat_region		/rpt_family="MERSA"	
repeat_region		437..581	
repeat_region		/rpt_family="MERSA"	
repeat_region		complement(720..1035)	
repeat_region		/rpt_family="Aluub"	
repeat_region		complement(1052..1303)	
repeat_region		/rpt_family="AluSx"	
repeat_region		1354..1480	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(1492..1538)	
repeat_region		/rpt_family="AluSx"	
repeat_region		1641..1909	
repeat_region		/rpt_family="AluSx"	
repeat_region		complement(2011..2123)	
repeat_region		/rpt_family="MERSA"	
repeat_region		2124..2427	
repeat_region		/rpt_family="AluSc"	
repeat_region		complement(2428..2496)	
repeat_region		/rpt_family="MERSA"	
repeat_region		complement(2600..2643)	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(2644..2952)	
repeat_region		/rpt_family="Aluub"	
repeat_region		complement(2953..3098)	
repeat_region		/rpt_family="MIR"	
repeat_region		3175..3455	
repeat_region		/rpt_family="AluSx"	
repeat_region		complement(4937..5019)	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(7705..7809)	
repeat_region		/rpt_family="MIR"	
repeat_region		8533..8603	
repeat_region		/rpt_family="(TCTG)n"	
repeat_region		10063..10265	
repeat_region		/rpt_family="G-rich"	
repeat_region		11845..11885	
repeat_region		/rpt_family="(CAGAGA)n"	
repeat_region		11911..12081	
repeat_region		/rpt_family="(GA-rich)"	
repeat_region		complement(12936..13004)	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(13094..13417)	
repeat_region		/rpt_family="Aluub"	
repeat_region		complement(14200..14371)	
repeat_region		/rpt_family="MIR"	
repeat_region		complement(14460..14771)	
repeat_region		/rpt_family="AluSx"	
repeat_region		15587..15717	
repeat_region		/rpt_family="(CAGAGA)n"	
repeat_region		15621..15721	
repeat_region		/rpt_family="(GA-rich)"	
repeat_region		complement(16362..16472)	
repeat_region		/rpt_family="L2"	
repeat_region		17700..17754	
repeat_region		/rpt_family="(GC-rich)"	
repeat_region		18133..18162	

repeat_region		/rpt_family="(GC-rich)"	
repeat_region		18191..18211	
repeat_region		/rpt_family="(GC-rich)"	
repeat_region		18337..18458	
repeat_region		/rpt_family="(CCCCG)n"	
repeat_region		18651..18810	
repeat_region		/rpt_family="MIR3"	
repeat_region		19403..19442	
repeat_region		/rpt_family="(TG)n"	
repeat_region		complement(20426..20751)	
repeat_region		/rpt_family="AluSx"	
repeat_region		21241..21279	
repeat_region		/rpt_family="(TCCC)n"	
repeat_region		21752..21960	
repeat_region		/rpt_family="Aluuo"	
repeat_region		complement(22195..22501)	
repeat_region		/rpt_family="Aluv"	
repeat_region		complement(23668..23962)	
repeat_region		/rpt_family="AluSp"	
repeat_region		24954..25095	
repeat_region		/rpt_family="MER115"	
repeat_region		25126..25422	
repeat_region		/rpt_family="AluSx"	
repeat_region		complement(25496..25604)	
repeat_region		/rpt_family="FLAM_A"	
repeat_region		25809..25834	
repeat_region		/rpt_family="(CA)n"	
repeat_region		complement(25851..26165)	

Query Match		96.7%; Score 6032; DB 9; Length 122132;	
Best local similarity		99.2%; Pred. No. 0;	
Matches 6190; Conservative		0; Mismatches 35; Indels 17; Gaps 12;	
QY	1	GATCACTTGAGGACAGTATGTCACAGCCAGCCCTGGGCGCATAGGAGACTGTCTCTACG	60
DB	72289	GATCACTTGAGGACAGTATGTCACAGCCAGCCCTGGGCGCATAGGAGACTGTCTCTACG	72230
QY	61	AAAAATCAAAAAATTATGCGCGGCGCATGTGCTCAGCTCTGTAATCCCTGAACCTTTGGG	120
DB	72229	AAAAATCAAAAAATTATGCGCGGCGCATGTGCTCAGCTCTGTAATCCCTGAACCTTTGGG	722170
QY	121	ACATCAAGGCAAGTGGATCAGTGGATGAGAGTGGAGACTAGCCGGGCAACATGCT	180
DB	72169	ACATCAAGGCAAGTGGATCAGTGGATGAGAGTGGAGACTAGCCGGGCAACATGCT	722110
QY	181	GAACCCCTATCTCCATAAAAATACAAAAATTAGCCAGGATGGTGCGACGACCTGTA	240
DB	72109	GAACCCCTATCTCTACT-AAAAATACAAAAATTAGCCAGGATGGTGCGACGACCTGTA	72051
QY	241	ATCCCGGCTACTCAGGAGGCTGAGGCGAGAGATCACTTGAACCCAGGAGCGGAGTTG	300
DB	72050	ATCCCGGCTACTCAGGAGGCTGAGGCGAGAGATCACTTGAACCCAGGAGCGGAGTTG	71991
QY	301	CAGTGGCTGAGTACACACCTGACCTCCAGCCCTGGGTGACAGACAGACCTCTATCTC	360
DB	71990	CAGTGGCTGAGTACACACCTGACCTCCAGCCCTGGGTGACAGACAGACCTCTATCTC	71931
QY	361	AAAAAAATATAAAAAATTAATAATTAAGCCAGGATGGTATGACACCTTAGTCTCAG	420
DB	71930	AAAAAAATATAAAAAATTAATAATTAAGCCAGGATGGTATGACACCTTAGTCTCAG	71871
QY	421	CTACTCAGGAGGCTGAGGTGGGAGATCACTTGAACCTGGGCGAGTCAAGGCTCAAGTGA	480
DB	71870	CTACTCAGGAGGCTGAGGTGGGAGATCACTTGAACCTGGGCGAGTCAAGGCTCAAGTGA	71812
QY	481	GCCAAAGTACGACGACATCACTCAGCTGGGCAACAGAGAGAACCTGTCTTAAAAA	540
DB	71811	GCCAAAGTACGACGACATCACTCAGCTGGGCAACAGAGAGAACCTGTCTTAAAAA	71752
QY	541	AATAATATAATATAAGAAAAAAGAGCTCTGTTATGCTCTGTCATACATACACT	600
DB	71751	AATAATATAATATAAGAAAAAAGAGCTCTGTTATGCTCTGTCATACATACACT	71695

QY 601 ATGTATATAGTTTGCACAACTCAAGATCCAGATAGTCATTAATTTTTAGGCTGTGGCCGT 660
|||||
Db 71694 ATGTATATAGTTTGCACAACTCAAGATCCAGATAGTCATTAATTTTTAGGCTGTGGCCGT 71635
QY 661 ATGGTCTGTGTCAATCACTGTGCGCTGTCTTCTTTCAGCAAAAGAGAGCTATTAACAT 720
|||||
Db 71634 ATGGTCTGTGTCAATCACTGTGCGCTGTCTTCTTTCAGCAAAAGAGAGCTATTAACAT 71575
QY 721 ACATACATGAATTTTATATAGATCAGATTTGAATTTATATATATATATATATATAT 780
|||||
Db 71574 ACATACATGAATTTTATATAGATCAGATTTGAATTTATATATATATATATATATAT 71515
QY 781 AAAATATCTTTTAAAAATTTTCCCTAACATTTAAAAAGTAAAAAGCCGCA--G 837
|||||
Db 71514 AAAATATCTTTTAAAAATTTTCCCTAACATTTAAAAAGTAAAAAGCCGCAAGCG 71455
QY 838 CGGGCCATGCTCACGCGCTGTATTTCCAGCACTTTGGAGGCTGAGTGGGAGATCACTT 897
|||||
Db 71454 CGGGCCATGCTCACGCGCTGTATTTCCAGCACTTTGGAGGCTGAGTGGGAGATCACTT 71395
QY 898 GAGATCAACAGTTTCAGAGACCGCTGGCCAAATAGCAAAACCCCATTTCTACTAAAAAT 957
|||||
Db 71394 GAGATCAACAGTTTCAGAGACCGCTGGCCAAATAGCAAAACCCCATTTCTACTAAAAAT 71335
QY 958 AAAAAATTAAGTGGGCATAGTGGTGACACCTGTGTATCCAGCTACTTGGAGGCTGAG 1017
|||||
Db 71334 AAAAAATTAAGTGGGCATAGTGGTGACACCTGTGTATCCAGCTACTTGGAGGCTGAG 71275
QY 1018 GCAGGAAATCGCTTGAACTGGGAAAGGAGGTGAGTGCAGCAATCATATCCCATG 1077
|||||
Db 71274 GCAGGAAATCGCTTGAACTGGGAAAGGAGGTGAGTGCAGCAATCATATCCCATG 71215
QY 1078 CACTCCAGCGCTGGGTGACAGAGTGAAGTCTGTCTCAACGAAAAAAGTGTAAAGC 1137
|||||
Db 71214 CACTCCAGCGCTGGGTGACAGAGTGAAGTCTGTCTCAACGAAAAAAGTGTAAAGC 71155
QY 1138 CATTCCTAATTCAGTGTACATCACTAGTGTACTAGTCTGCTGCTGAGG 1197
|||||
Db 71154 CATTCCTAATTCAGTGTACATCACTAGTGTACTAGTCTGCTGCTGAGG 71095
QY 1198 CATACCTGAGACAGTGAAGTGTGTGTGCACAGGACATACATTTCCACATTAACATGAC 1257
|||||
Db 71094 CATACCTGAGACAGTGAAGTGTGTGTGCACAGGACATACATTTCCACATTAACATGAC 71035
QY 1258 ACTACCAAGTTGCATCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1317
|||||
Db 71034 ACTACCAAGTTGCATCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 70975
QY 1318 AATGAGAGTTACTTCAGATCTCTTTACAAAGATGCTTAAGCCCAAGTACCAAGATGAACA 1377
|||||
Db 70974 AATGAGAGTTACTTCAGATCTCTTTACAAAGATGCTTAAGCCCAAGTACCAAGATGAACA 70915
QY 1378 GGAAGTGGAGGGGAAAGCTCCAGCCCTTCTAACCATGTAAGAAATCTGGTGAAGCCT 1437
|||||
Db 70914 GGAAGTGGAGGGGAAAGCTCCAGCCCTTCTAACCATGTAAGAAATCTGGTGAAGCCT 70855
QY 1438 TCTGGATGCTGAAGAGTGAATAACGGGGTCTCTGAGGCTGCGCCCTGTCAAGTACT 1497
|||||
Db 70854 TCTGGATGCTGAAGAGTGAATAACGGGGTCTCTGAGGCTGCGCCCTGTCAAGTACT 70795
QY 1498 GTGACTTCTAGCGCTCCAGTCAGATCTCAGCCCCCATGTGTCATGCGCAGTATATAGC 1557
|||||
Db 70794 GTGACTTCTAGCGCTCCAGTCAGATCTCAGCCCCCATGTGTCATGCGCAGTATATAGC 70736
QY 1558 CCTCAGCTCTGTGTTGGTCTTTATCTC-CCCATGTGGGGCTGAAGCTGGATGAACCG 1616
|||||
Db 70735 CCTCAGCTCTGTGTTGGTCTTTATCTCCCCCATGTGGGGCTGAAGCTGGATGAACCG 70676
QY 1617 TTAATCAAGATGTACAGCTTCTTGAACGAAAGTATGTACAGAAACAGACAGGGCTT 1676
|||||
Db 70675 TTAATCAAGATGTACAGCTTCTTGAACGAAAGTATGTACAGAAACAGACAGGGCTT 70616
QY 1677 GGCAGAGTATGTACTACCAATCTCACTGGCTCAGCCACAGCTAGTTCTGTGATCTT 1736

Db 70615 GGCAGAGTATGTACTACCAAAATCCTACCGGGCTCAGCCACAGCTAGTTCTGTGATCTT 70556
QY 1737 GAACAAATTTTCTACTCTCTGAGGCAATCCCTGTGCTCAACACACAGAGTGGTTGAC 1796
|||||
Db 70555 GAACAAATTTTCTACTCTCTGAGGCAATCCCTGTGCTCAACACACAGAGTGGTTGAC 70496
QY 1797 AGATGAATAGAGC-AAGTCCCTTACACTGTATCCACACTTTTGGAGGCCAAGCGC 1855
|||||
Db 70495 AGATGAATAGAGCAAAAGTGCTTACACCTGTATCCACACTTTTGGAGGCCAAGCGC 70436
QY 1856 GGTGGATGCTTGAGCCTGAGAGTGAACACATATCCGGCAGTCTCAGACCCCTGTTG 1915
|||||
Db 70435 GGTGGATGCTTGAGCCTGAGAGTGAACACATATCCGGCAGTCTCAGACCCCTGTTG 70376
QY 1916 CTCTGGGCGCTCCTCTGCGCTGGGCTCCCACTTCCGTGGACATTTAGAGACCCCTTCA 1975
|||||
Db 70375 CTCTGGGCGCTCCTCTGCGCTGGGCTCCCACTTCCGTGGACATTTAGAGACCCCTTCA 70316
QY 1976 CACCGCTGCACTGTGGAGGCCCTTTCTGGGCTGGCCAAAGCCAGAGCCGGCTCCAG 2035
|||||
Db 70315 CACCGCTGCACTGTGGAGGCCCTTTCTGGGCTGGCCAAAGCCAGAGCCGGCTCCAG 70256
QY 2036 CTTCAGAGGAGGTGTGGAGGAGAGGCTCAAGCAGAACCGGGGCTGCGACGGCGCTTG 2095
|||||
Db 70255 CTTCAGAGGAGGTGTGGAGGAGAGGCTCAAGCAGAACCGGGGCTGCGCGCGCTTG 70196
QY 2096 CGGGCCAGCTGTGAGTTCCGGGTGGGCGTGGGCTGGGCGGCCCGCACTGGAGCAGCG 2155
|||||
Db 70195 CGGGCCAGCTGTGAGTTCCGGGTGGGCGTGGGCTGGGCGGCCCGCACTGGAGCAGCG 70136
QY 2156 GCCAGCCCTCCAGGCCCCCGGCAATGAGAGGCTTACACCCGGGCGAGCGGCTGCGAG 2215
|||||
Db 70135 GCCAGCCCTCCAGGCCCCCGGCAATGAGAGGCTTACACCCGGGCGAGCGGCTGCGAG 70076
QY 2216 GGTGTACTGGGTGCCCAAGATGCCAGCCCGCGCGCTGTGTCTGCTGATTTCTAC 2275
|||||
Db 70075 GGTGTACTGGGTGCCCAAGATGCCAGCCCGCGCGCTGTGTCTGCTGATTTCTAC 70016
QY 2276 TGGGCTTACAGGCTTCCCGCGGGGAGGAGGCTGCGGACCTGCGCCCATACCTGAG 2335
|||||
Db 70015 TGGGCTTACAGGCTTCCCGCGGGGAGGAGGCTGCGGACCTGCGCCCATACCTGAG 69956
QY 2336 CTTCCCTCATGGGCTCTGTGCGGCGAGCCTCCCGACAGACACACCCCTGCTC 2395
|||||
Db 69955 CTTCCCTCATGGGCTCTGTGCGGCGAGCCTCCCGACAGACACACCCCTGCTC 69896
QY 2396 CACAGCGCCCAATGCCATTCAGCACGCAAGAGGCTGAGAAATGGGGGCGCACGGACGG 2455
|||||
Db 69895 CACAGCGCCCAATGCCATTCAGCACGCAAGAGGCTGAGAAATGGGGGCGCACGGACGG 69836
QY 2456 ACTGGCAGGAGCTACCCCTGCAAGCCCTGTGGGAATCCACTGGGTGAAGCAGCTGGG 2515
|||||
Db 69835 ACTGGCAGGAGCTACCCCTGCAAGCCCTGTGGGAATCCACTGGGTGAAGCAGCTGGG 69776
QY 2516 CTCTGAGTGTGTGAGACTTGGAGAACTTTATGTCTAGCTCAGGGATGTAATATCA 2575
|||||
Db 69775 CTCTGAGTGTGTGAGACTTGGAGAACTTTATGTCTAGCTCAGGGATGTAATATCA 69716
QY 2576 CCAATCAGCACCCGTGTCTAGCTCAGAGGCTGTGAAATGCACCAATCAACATCTGTATC 2635
|||||
Db 69715 CCAATCAGCACCCGTGTCTAGCTCAGAGGCTGTGAAATGCACCAATCAACATCTGTATC 69656
QY 2636 TAGTACTCTGATGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAATATCAC 2695
|||||
Db 69655 TAGTACTCTGATGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAATATCAC 69596
QY 2696 CAATCGCAGCTGTATCTAGCTCAAGTGTGTAACACACCAATCAGACCCCTGTGTCT 2755
|||||
Db 69595 CAATCGCAGCTGTATCTAGCTCAAGTGTGTAACACACCAATCAGACCCCTGTGTCT 69536
QY 2756 AGCTCAGGGTATGTAATGCACCAATGACAGTGTGTATGTGGCTACTTTCATGGGATC 2815
|||||

Db 69335 AGCTCAGGGTATGTGAATGCACCAATCGACAGTCTGTATCTGGCTACTTTCATGGGCATC 69476
Oy 2816 CGTGTGAAGAGACCAACCAAGGCTTGTGTGAGCCATTAAGCTTCTATCATCCCGGTG 2875
Db 69475 CGTGTGAAGAGACCAACCAAGGCTTGTGTGAGCAATTAAGGCTTCTATCATCCCGGTG 69416
Oy 2876 CAGGTGGGCTGAGTCCGAAAAGAGTACGAGGAGGAGATTAAGGGTGGGCGCTTTAT 2935
Db 69415 CAGGTGGGCTGAGTCCGAAAAGAGTACGAGGAGGAGATTAAGGGTGGGCGCTTTAT 69336
Oy 2936 AGAATTTGGGTAGGTAAAGAAAATTTACATGTCAAAGGGGTTTGTCTGTGGCGGGCAGG 2995
Db 69335 AGAATTTGGGTAGGTAAAGAAAATTTACATGTCAAAGGGGTTTGTCTGTGGCGGGCAGG 69296
Oy 2996 AG-TGGGGGGTCCGAAAGGTGCTCAGTGGGGGTGCTTTTGAAGCCAGGATAGCCGAGAA 3054
Db 69295 AGTTGGGGGGTCCGAAAGGTGCTCAGTGGGGGTGCTTTTGAAGCCAGGATAGCCGAGAA 69236
Oy 3055 AGGACTTTCACAAAGGTATGTCATCAATTAAGGCAAGGACCCGCAATTTACACCTTTT 3114
Db 69235 AGGACTTTCACAAAGGTATGTCATCAATTAAGGCAAGGACCCGCAATTTACACCTTTT 69176
Oy 3115 GTGGTGAATGTCATCACTTAAGTTGGGGCAGGGCATATTCATCTTTTGTGATTCCTC 3174
Db 69175 GTGGTGAATGTCATCACTTAAGTTGGGGCAGGGCATATTCATCTTTTGTGATTCCTC 69116
Oy 3175 AGTTACTTCAGGCACTGGGGGTATATGTCAAGTTACAGGGATAGGATGGCTTGGCT 3234
Db 69115 AGTTACTTCAGGCACTGGGGGTATATGTCAAGTTACAGGGATAGGATGGCTTGGCT 69056
Oy 3235 TGGGCTCAGAGGCTTGAGAGTACTCTGTGGGGGCTTGAGAAATGTTTGTGTGCACACT 3294
Db 69055 TGGGCTCAGAGGCTTGAGAGTACTCTGTGGGGGCTTGAGAAATGTTTGTGTGCACACT 68996
Oy 3295 CTGTATCTAGTTAATCTAGTGGGGCAGTGGAAACCTTTGTGTCTAGTCTCAGGAGTTGTA 3354
Db 68995 CTGTATCTAGTTAATCTAGTGGGGCAGTGGAAACCTTTGTGTCTAGTCTCAGGAGTTGTA 68936
Oy 3355 AACGCAACCAATCAGGGCCCTGCACAAACAGACCACTGGGCTCTACCAATCAGCAGAGATG 3414
Db 68935 AACGCAACCAATCAGGGCCCTGCACAAACAGACCACTGGGCTCTACCAATCAGCAGAGATG 68876
Oy 3415 GGGTGGGGCAGATTAAGAAATTAAGAGCAGGCTGCCAGCAGCAGAGTGGCAACCGCAC 3474
Db 68875 GGGTGGGGCAGATTAAGAAATTAAGAGCAGGCTGCCAGCAGCAGAGTGGCAACCGCAC 68816
Oy 3475 AGGTCCCTATCCACAAATATGSCAGCTTTGTTTGTGTGTGGATAAATCTTGTCTAC 3534
Db 68815 AGGTCCCTATCCACAAATATGSCAGCTTTGTTTGTGTGTGGATAAATCTTGTCTAC 68756
Oy 3535 TGTCTGCTTTTGGGTCCACACTGCTTTTATGAGCTGTAACTATACACGAGGTCTGC 3594
Db 68755 TGTCTGCTTTTGGGTCCACACTGCTTTTATGAGCTGTAACTATACACGAGGTCTGC 68696
Oy 3595 AGCTTCACTCTGGAAGCCATTAAGACCAAGCAGGCCAGGGGAGGATGAACAACTCCGGC 3654
Db 68695 AGCTTCACTCTGGAAGCCATTAAGACCAAGCAGGCCAGGGGAGGATGAACAACTCCGGC 68636
Oy 3655 CGGCTGCTTAAAGAGCTATAACACTCCCGGAGAGTCTGTGAGCTTCTACTCTCCAGCCA 3714
Db 68635 CGGCTGCTTAAAGAGCTATAACACTCCCGGAGAGTCTGTGAGCTTCTACTCTCCAGCCA 68576
Oy 3715 GCGAGACCAAGACCCACCAAGAGAGAGAACTGGCAACATCTGAACATCAGAGGAA 3774
Db 68575 GCGAGACCAAGACCCACCAAGAGAGAGAACTGGCAACATCTGAACATCAGAGGAA 68516
Oy 3775 CAAACTCCAGATGACCACTTAAAGAGCTTAACACTAGCGGAGGGTCCGGGCTTCC 3834
Db 68515 CAAACTCCAGATGACCACTTAAAGAGCTTAACACTAGCGGAGGGTCCGGGCTTCC 68456
Oy 3835 TTCTTGAAGTCAAGTCAAGCACTCACCAGTTTGGAGACACAAGCCAGAGTTTGA 3894
Db 68455 TTCTTGAAGTCAAGTCAAGCACTCACCAGTTTGGAGACACAAGCCAGAGTTTGA 68336

Oy 3895 ATCAGCGTGGGCAACATGATGAATGCCCTCTGCAAAAAAAAAAAAAAAAAATTACAAAAT 3954
Db 68395 ATCAGCGTGGGCAACATGATGAATGCCCTCTGCAAAAAAAAAAAAAAAAAATTACAAAAT 68336
Oy 3955 TGGCGAGACATGATGATCCGTGCTGTGGTCCACACTACGGGGAGGCTTAAAGTGGAGG 4014
Db 68335 TGGCGAGACATGATGATCCGTGCTGTGGTCCACACTACGGGGAGGCTTAAAGTGGAGG 68276
Oy 4015 ATCGCTTGAACCTGGAGGTGAAGACTGCACTGATGATTTGAACACAGCCCTTAG 4074
Db 68275 ATCGCTTGAACCTGGAGGTGAAGACTGCACTGATGATTTGAACACAGCCCTTAG 68216
Oy 4075 GCTGGGGGACAGTGAAGACCTGTTTCCCTCCGCAAAAAAAAAATTGACAAAAGTGAATA 4134
Db 68215 GCTGGGGGACAGTGAAGACCTGTTTCCCTCCGCAAAAAAAAAATTGACAAAAGTGAATA 68136
Oy 4135 AGAGTGCCTGATATGCTAGAGCGCAGTGCATGCTGTAAATCCACAGCACTTTGGGA 4194
Db 68135 AGAGTGCCTGATATGCTAGAGCGCAGTGCATGCTGTAAATCCACAGCACTTTGGGA 68096
Oy 4195 GCCGAGCGGGCGGCTACCTTAAGTCAAGATGTGAACACAGCCTGGCCAAATGAGGA 4254
Db 68095 GCCGAGCGGGCGGCTACCTTAAGTCAAGATGTGAACACAGCCTGGCCAAATGAGGA 68036
Oy 4255 AAGCCATCTCTTCAAAAAATCAAAAAATTAAGCGGGCTGTGGGGCAGTGTGAGCATGC 4314
Db 68035 AAGCCATCTCTTCAAAAAATCAAAAAATTAAGCGGGCTGTGGGGCAGTGTGAGCATGC 67976
Oy 4315 CTGTATCCACACTCTCAGAGGCTGAGAGCAAGAAATCACTTGAACCCAGAGAGCGGC 4374
Db 67975 CTGTATCCACACTCTCAGAGGCTGAGAGCAAGAAATCACTTGAACCCAGAGAGCGGC 67916
Oy 4375 GGTTCAGTGAAGCCAGATCGTCCATGTCACCTCAACCATCTCCAGCTTGGCAACAAGA 4434
Db 67915 GGTTCAGTGAAGCCAGATCGTCCATGTCACCTCAACCATCTCCAGCTTGGCAACAAGA 67856
Oy 4435 GCCAAACCTGTGCTT - AAAAAAAAAAAAAAAAAAAGTCCGTGCAATTAAGAGTGTGCA 4492
Db 67855 GCCAAACCTGTGCTT AAAAAAAAAAAAAAAAAAAGTCCGTGCAATTAAGAGTGTGCA 67796
Oy 4493 TGCATATGTTGCAGGCAACATGTTTAAAGATGTGAGAGTCTCGCTTCCATGATGCTGT 4552
Db 67795 TGCATATGTTGCAGGCAACATGTTTAAAGATGTGAGAGTCTCGCTTCCATGATGCTGT 67736
Oy 4553 TAAAAACCCACCTCAAGGCCAGGTGCAAGTGCATGCTCAATTAATCCAGCATTTGGG 4612
Db 67735 TAAAAACCCACCTCAAGGCCAGGTGCAAGTGCATGCTCAATTAATCCAGCATTTGGG 67676
Oy 4613 AGCGGAGCGGGTGGATCACCCTGAGGTCAAGAGTTCCAGACCACTGACCAACAT 4672
Db 67675 AGCGGAGCGGGTGGATCACCCTGAGGTCAAGAGTTCCAGACCACTGACCAACAT 67616
Oy 4673 GGTGAATCCCACTCTACTTAAATTAACAAATTAAGATGAGCATGGTGTGATGCTGT 4732
Db 67615 GGTGAATCCCACTCTACTTAAATTAACAAATTAAGATGAGCATGGTGTGATGCTGT 67556
Oy 4733 AATCCACCTACTTGGAGGCTGAGCAGGAAATCACTAGAACCAAGGAGCGGAGTT 4792
Db 67555 AATCCACCTACTTGGAGGCTGAGCAGGAAATCACTAGAACCAAGGAGCGGAGTT 67496
Oy 4793 GTAAGTGAAGATGCTGCTTGTCACTCCAGCTGAGCAATGAGCAACTCCATCTC 4852
Db 67495 GTAAGTGAAGATGCTGCTTGTCACTCCAGCTGAGCAATGAGCAACTCCATCTC 67436
Oy 4853 AAAAAAACCAACAAAAAACCAACCTCTACTCCAGGAGGCTGGTACAGAGTGGGCG 4912
Db 67435 AAAAAAACCAACAAAAAACCAACCTCTACTCCAGGAGGCTGGTACAGAGTGGGCG 67376
Oy 4913 ACATAGTGAAGGCTGAGCCACAGAGCTTAAGCGGAGGCTGACAGACCGGAGCAGGA 4972
Db 67375 ACATAGTGAAGGCTGAGCCACAGAGCTTAAGCGGAGGCTGACAGACCGGAGCAGGA 67316

QY 4973 TAAACAGTGTGAGATCATGTGTGAGATCAGACGTCCCTGCCATTTGGTGCACACAGGG 5032
|||||
Db 67315 TAAACAGTGTGAGATCATGTGTGAGATCAGACGTCCCTGCCATTTGGTGCACACAGGG 67256
QY 5033 GGCCCCCAGACACAGAGATGGCCCCATCCAGTCCACACATCCACTTCTCATCCAGAGAT 5092
|||||
Db 67255 GGCCCCCAGACACAGAGATGGCCCCATCCAGTCCACACATCCACTTCTCATCCAGAGAT 67196
QY 5093 GTCTGTTCTTGGGACCTGGGGGTAAATTAGACAGAAAGTGACAGTCTTGGGTGTGC 5152
|||||
Db 67195 GTCTGTTCTTGGGACCTGGGGGTAAATTAGACAGAAAGTGACAGTCTTGGGTGTGC 67136
QY 5153 AGTCAGACTGCCCCAGGACAGGCTTGTGGCTGTAGAAAACGTTCCAGGCCCTAGGCCGGGC 5212
|||||
Db 67135 AGTCAGACTGCCCCAGGACAGGCTTGTGGCTGTAGAAAACGTTCCAGGCCCTAGGCCGGGC 67076
QY 5213 ACGGTGGCTCAGCGCTTAATCCACAGACTTTTGGAGAGCGAGGCGGGTGGATCAGAGG 5272
|||||
Db 67075 ACGGTGGCTCAGCGCTTAATCCACAGACTTTTGGAGAGCGAGGCGGGTGGATCAGAGG 67016
QY 5273 TCAGAGATCGTGACATCTGCTAGCTAACACGGTGAAACCCCGTCTCTACTAAAAATCAA 5332
|||||
Db 67015 TCAGAGATCGTGACATCTGCTAGCTAACACGGTGAAACCCCGTCTCTACTAAAAATCAA 66956
QY 5333 AAAATTTGGCGGGGAGTGGCGGGACCTGTAGTTCAGCTACTCGGGAGCGTACAGCA 5392
|||||
Db 66955 AAAATTTGGCGGGGAGTGGCGGGACCTGTAGTTCAGCTACTCGGGAGCGTACAGCA 66896
QY 5393 GGAGAAATGGCGTGAACCCGAGAGCAAGTTTGGAGTACAGCCGAGTCCGCCACTGCAC 5452
|||||
Db 66895 GGAGAAATGGCGTGAACCCGAGAGCAAGTTTGGAGTACAGCCGAGTCCGCCACTGCAC 66836
QY 5453 TCCAGCTGGCGCAGACAGCAAGACTCTGTGAAAAAGAAAAGAAAGCTTCAGGTCT 5512
|||||
Db 66835 TCCAGCTGGCGCAGACAGCAAGACTCTGTGAAAAAGAAAAGAAAGCTTCAGGTCT 66776
QY 5513 GAGCAGAGGCGCCAGGCTTAATCTGTCACTTACATGACCTGGGCAAGGCATCTCCT 5572
|||||
Db 66775 GAGCAGAGGCGCCAGGCTTAATCTGTCACTTACATGACCTGGGCAAGGCATCTCCT 66716
QY 5573 TCCGTGGCCAGTTACAGGGGTTGGAAATCGACTCAAGGTCCCTCCAGCATTTAAAGCTG 5632
|||||
Db 66715 TCCGTGGCCAGTTACAGGGGTTGGAAATCGACTCAAGGTCCCTCCAGCATTTAAAGCTG 66656
QY 5633 CATGGTTCTTAAGATGAGAAATGGGGCAAGTTTCCCTCTCACCCAGCCCGTGTCCAC 5692
|||||
Db 66655 CATGGTTCTTAAGATGAGAAATGGGGCAAGTTTCCCTCTCACCCAGCCCGTGTCCAC 66596
QY 5693 TTCAAGGTGAATGACAGGGGAGTCAAGTGTCCCAATCCCGCAAGTCCAAAGCCCTGGG 5752
|||||
Db 66595 TTCAAGGTGAATGACAGGGGAGTCAAGTGTCCCAATCCCGCAAGTCCAAAGCCCTGGG 66536
QY 5753 GACCCTACTGTACAGGGTCTGTCAGACGAGAGGTGAAGGTCAAGTGAAGCAATCCGCTGAA 5812
|||||
Db 66535 GACCCTACTGTACAGGGTCTGTCAGACGAGAGGTGAAGGTCAAGTGAAGCAATCCGCTGAA 66476
QY 5813 GGCTCTTGCTCATTCGGGACAGACATCCGGTTTCTGTGGCTTACCCGGGATTTAAGG 5872
|||||
Db 66475 GGCTCTTGCTCATTCGGGACAGACATCCGGTTTCTGTGGCTTACCCGGGATTTAAGG 66416
QY 5873 GCTTTAGCCGAATGAGTCAATGGGGGGGGGGGGTCTGTGGGAGTTCCTCCAGTATCA 5932
|||||
Db 66415 GCTTTAGCCGAATGAGTCAATGGGGGGGGGGGGTCTGTGGGAGTTCCTCCAGTATCA 66356
QY 5933 ACTTTGGACAGACAGCTGGAATCTTGCATGTGCTTATCCAAAGTGTGGGGTGGGACA 5992
|||||
Db 66355 ACTTTGGACAGACAGCTGGAATCTTGCATGTGCTTATCCAAAGTGTGGGGTGGGACA 66296
QY 5993 GCAAGCCAAAGCCCAATGCTTATCTCAGGTAGGGGCTCAGAGAGTCTCCAGACAGCA 6052
|||||
Db 66295 GCAAGCCAAAGCCCAATGCTTATCTCAGGTAGGGGCTCAGAGAGTCTCCAGACAGCA 66236
QY 6053 GCCTCCGGAAGTTTGGGGGTAGGAATGGAGCAACCAAGCCTCTTTTCTCTCTTAG 6112

Db 66235 GCCTCCGAGAGTTTGGGGGTAGGAATGGAGCAACA-GCTTTTCTTTCTCTCTTAG 66177
QY 6113 AATTTGGGGGCTTGGGGACAGGCTTGAATCCCAAGAGAGAGGCGCAAGGACACT-C 6171
|||||
Db 66176 AATTTGGGGGCTTGGGGACAGGCTTGAATCCCAAGAGAGAGGCGCAAGGACACTGC 66117
QY 6172 CCCCACAAGTGTGCCAGACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6231
|||||
Db 66116 CCCCACAAGTGTGCCAGAGC-AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 66058
QY 6232 GC 6233
||
Db 66057 GC 66056

RESULT 7
AC007776
LOCUS
DEFINITION Homo sapiens chromosome 17 clone hC17.413_A4 map 17, ***
AC007776
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 100490)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone hC17.413_A4
Unpublished
2 (bases 1 to 100490)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckler, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeRubeis, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heath, A., Horton, L., Howland, J., Jones, C., Kann, L.,
Karatas, A., Lebeck, J., Liu, C., Locke, K., MacDonald, P.,
Margolis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nilot, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Teeffaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
Direct Submission
Submitted (10-JUN-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2868: contig of 2868 bp in length
*
*
* 2869 6849: contig of 3981 bp in length
*
*
* 6850 11677: contig of 4828 bp in length
*
*
* 11678 16813: contig of 5136 bp in length
*
*
* 16814 23727: contig of 6914 bp in length
*
*
* 23728 37142: contig of 13415 bp in length

* gap of unknown length
* 37143 49015: contig of 11873 bp in length
* 49016 62202: contig of 13187 bp in length
* 62203 79211: contig of 17009 bp in length
* 79212 100490: contig of 21279 bp in length.
* Location/Qualifiers
 1. 100490
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /clone="hct.413_A_4"
 /clone_lib="Research Genetics/cal Tech CITB-HSP-C (plates 195-384)"
BASE COUNT 24217 a 25467 c 26408 g 23679 t 719 others
ORIGIN
Query Match 91.0%: Score 5671.6; DB 2: Length 100490;
Best Local Similarity 97.98; Pred. No. 0;
Matches 6153: Conservative 0; Mismatches 71; Indels 62; Gaps 39;
QY 2 ATCACTTGAGAGCAGTAGTTCACAGACGAGCCTGGGACACATAGGAGAGCTGTCTACGA 61
Db 63208 ATCACTTGAGAGCAGTAGTTCACAGACGAGCCTGGGACACATAGGAGAGCTGTCTACGA 63267
QY 62 AAAATCAAAAAAATATATGCGCGGCGATGGTGGCTCAGCTCTGTAACTCCCTGAACCTTGGGA 121
Db 63268 AAAATCAAAAAAATATGCGCGGCGATGGTGGCTCAGCTCTGTAACTCCCTGAACCTTGGGA 63347
QY 122 CATCAAGCAGAGTGGATCTACTTGAAGTTCAGAGATTCAGACTAGCCTGGCCAACTGGTG 181
Db 63328 CATCAAGCAGAGTGGATCTACTTGAAGTTCAGAGATTCAGACTAGCCTGGCCAACTGGTG 63387
QY 182 AAACCTATCTCCACTAAAAAATATACAAAAATTAAGCCAGGATGGTGGCGAGCCTGTAA 241
Db 63388 AAACCTATCTCTACT-AAAAATACAAAAATTAAGCCAGGATGGTGGCGAGCCTGTAA 63446
QY 242 TCCCGGCTACTCAGAGGCTGAGCGAGGATCACTTGAACCCAGAGGCGGAGGTTGC 301
Db 63447 TCCCGGCTACTCAGAGGCTGAGCGAGGATCACTTGAACCCAGAGGCGGAGGTTGC 63506
QY 302 AGTAGGCTGAGATCAACCACTGCATCCAGCTGGGGTGGAGACAGAGCACTCTATCTCA 361
Db 63507 AGTAGGCTGAGATCAACCACTGCATCCAGCTGGGGTGGAGACAGAGCACTCTATCTCA 63566
QY 362 AAAAAAATAAAAAATAAAAAATTAAGCCAGGATGGTGGCGAGCCTGTAGTCTAGC 421
Db 63567 AAAAAAATAAAAAATAAAAAATTAAGCCAGGATGGTGGCGAGCCTGTAGTCTAGC 63626
QY 422 TACTCAGAGGCTGAGTGGAGGATCACTTGAACCTGGGGCAGTCAAGGCTACAGTAG 481
Db 63627 TACTCAGAGGCTGAGTGGAGGATCACTTGAACCT- GGGCAGTCAAGGCTGAGTAG 63685
QY 482 CCAAGATCATCCCTACACTCCAGCCTGGGCAACAGAGAGACCTGTCT-CTAAAAA 540
Db 63686 CCAAGATCATCCCTACACTCCAGCCTGGGCAACAGAGAGACCTGTCTCTGCTAAAAA 63745
QY 541 AATAATATATATAAGAAAAAAGAGCTCTGTTTATGCTCTCTGCTCATATCAATCACT 600
Db 63746 AATAATATATATAAGAAAAAAGAGCTCTGTTTATGCTCTCTGCTCATATCA--TACT 63802
QY 601 ATGATATAGTTTCAAACTCAAAAGATCCAGATAGTCAATTTTAAAGCTGTGGGCGCT 660
Db 63803 ATGATATAGTTTCAAACTCAAAAGATCCAGATAGTCAATTTTAAAGCTGTGGGCGCT 63862
QY 661 ATGCTCTCTGCAATCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719
Db 63863 ATGCTCTCTGCAATCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63922

QY 720 TACATACATGAAATTTTATAGACATCGAGATTTGAATTTTATATATTTTACATTTTA 779
Db 63923 TACATACATGAAATTTTATAGACATCGAGATTTGAATTTTATATATTTTACATTTTA 63982
QY 780 TAAAAATATCTTTTAAAAATTTTCCCTTAACCATTTTAAAAAGTGTAAAGCGGCCA--- 836
Db 63983 TAAAAATATCTTTTAAAAATTTTCCCTTAACCATTTTAAAAAGTGTAAANAGCGGCCNCA 64042
QY 837 -----GGCGGCATCGTCAGCGCTGTATTTCCACACACTTTGGGAGGCTGAGGTGGCAG 890
Db 64043 GCGNCGCNCNATGGCTCAGCCCTGTATTTCCACACACTTTGGGAGGCTGAGGTGGCAG 64102
QY 891 ATCACTTGAGATCAACAGTTGAGAGACGAGCCCTGGCAACATATGCAAAACCCCATTTCTAC 950
Db 64103 ATCACTTGAGATCAACAGTTGAGAGACGAGCCCTGGCAACATATGCAAAACCCCATTTCTAC 64182
QY 951 TAAAAATAAAAAATTAAGCTGGGCAATAGTGTGACACACCTGTGATCCAGCTACTTGGGA 1010
Db 64183 TAAAAATAAAAAATTAAGCTGGGCAATAGTGTGACACACCTGTGATCCAGCTACTTGGGA 64222
QY 1011 GGCTGAGGCGAGAGATGCTTGAACCTGGGAGCGGAGGTTGCACTGAGCCACATCAT 1070
Db 64223 GGCTGAGGCGAGAGATGCTTGAACCTGGGAGCGGAGGTTGCACTGAGCCACATCAT 64282
QY 1071 GCCACTGACCTCAGCTGGGTGACA-GAGTGAGACTTCGTCTCAAG-AAAAAANAAG 1128
Db 64283 GCCACTGACCTCAGCTGGGTGACA-GAGTGAGACTTCGTCTCAAG-AAAAAANAAG 64342
QY 1129 TGTAAAGCCATTCCTTAATTCAGTGTACATCAGTGTACATCTAGCTCTGCTACTCT 1188
Db 64343 TGTAAAGCCATTCCTTAATTCAGTGTACATCAGTGTACATCTAGCTCTGCTACTCT 64402
QY 1189 GCTCTGAGGCACTCTGAGAGATGAGATGCTTGGTCAAGAGCATATACATTTCCACAT 1248
Db 64403 GCTCTGAGGCACTCTGAGAGATGAGATGCTTGGTCAAGAGCATATACATTTTCCACAT 64462
QY 1249 TAACTAGACACATCAAGATGGCATCCAGAGAG--TTTTTTTTTATACATCTACCTCC 1307
Db 64463 TAACTAGACACATCAAGATGGCATCCAGAGAGTTTTTTTTTATACATCTACCTCC 64522
QY 1308 CCCAGCAACAAATGAGAGTTACTCAGATCCTTTACAAAGATGCTTAAGCCAGATACA 1367
Db 64523 CCCAGCAACAAATGAGAGTTACTCAGATCCTTTACAAAGATGCTTAAGCCAGATACA 64582
QY 1368 GATGAAACAGAGAAATGGAGGAGGAGAGCTCCAGCC--CCTCTAACCATGAAATATCC 1426
Db 64583 GATGAAACAGAGAAATGGAGGAGGAGAGCTCCAGCCCTCTTAACCATGAAATATCC 64642
QY 1427 TGGTAGAGCCTTCTGATGCTGAGAGATGAAATACGGGGGCTCTGAGAGCTGCCCCCT 1486
Db 64643 TGGTAGAGCCTTCTGATGCTGAGAGATGAAATACGGGGGCTCTGAGAGCTGCCCCCT 64702
QY 1487 GTCAGATCACTGTGACTCTCTGAGCCTCCAGTCCAGTCTCAGGCCCATGTGTCTAGTGGCCAG 1546
Db 64703 GTCAGATCACTGTGACTCTCTGAGCCTCCAGTCCAGTCCAGTCTCAGGCCCATGTGTCTAGTGGCCAG 64761
QY 1547 TGAATATAGCCCTCACTCTCTGTTTGTCTTATTTCTCC--ATGTGGGGCTGAGTCT 1604
Db 64762 TGAATATAGCCCTCACTCTCTGTTTGTCTTATTTCTCCCATGATGAGGCTGAAATGC 64821
QY 1605 TGGATTTAGCCGCTTATTAAGATGTACAGCTTTCTTGACAGGAAGTGTGTACAGAAA 1664
Db 64822 TGGATTTAGCCGCTTATTAAGATGTACAGCTTTCTTGACAGGAAGTGTGTACAGAAA 64881
QY 1665 CAGAGGGGCTTGGCAAGATGATCTTAATGCAAAATCCATGCGGCTGAGCCACGCTAG 1724
Db 64882 CAGAGGGGCTTGGCAAGATGATCTTAATGCAAAATCCATGCGGCTGAGCCACGCTAG 64941
QY 1725 TTTCTGTATCTTGAACAAGTTTCTTCACTTCTGAGGCGCATCCCTTGGCTACAAACACAC 1784
Db 64942 TTTCTGTATCTTGAACAAGTTTCTTCACTTCTGAGGCGCATCCCTTGGCTACAAACACAC 65001
QY 1785 CAGTTG-GTTCAGAGAGATGAATAGAG-AAGTCCCTTACACTGTAAATCCAGACACTTGG 1842

D	65002	CAGTGTGNTTACAGAGATGAATGACGAAATGCTTACACTTATCTAATCCAGACTTTG	65061
Q	1843	GGAGGCCAAGGCGGGTGGATGGCTTGTGAGCCCTGAGAGGTGAGACAGCATCCGGCACTCTCA	1902
D	65062	GGAGGCCAAGGCGGGTGGATGGCTTGTGAGCCCTGAGAGGTGAGACAGCATCCGGCACTCTCA	65121
Q	1903	CAGCCCTCTGTTCGCTCTCGGCGCCTCTCTGCTTGGGCTCCCACTTGGTGGAAG	1962
D	65122	CAGCCCTCTGTTCGCTCTCGGCGCCTCTCTGCTTGGGCTCCCACTTGGTGGAAG	65181
Q	1963	GAGCCCTTACAGCCCAAGCCCTGCACTGTGTGGAGCCCTTCTGGGCTGGCCAGGCC--A	2019
D	65182	GAGCCCTTACAGCCCAAGCCCTGCACTGTGTGGAGCCCTTCTGGGCTGGCCAGGCCAGN	65241
Q	2020	GAGCCGGCTCCCTAGCTTGTGAGGG--AGGTGTGAGGAGAGAGGCTC--AAGCAAGACCG	2077
D	65242	AGCCGGCTCCCTAGCTTGTGAGGAGGAGGTGTGAGGAGGAGGCTCNAAGCAGGAGNACC	65301
Q	2078	GG--CTGGCAGCGGCGCTTGGCGGCAAGTGTGAGGCTCGGGTGGGCGTGGGCTT--GGCG	2133
D	65302	GGGCTGAGNCGGCGGCTTGGCGGCAAGTGTGAGGCTCGGGTGGGCGTGGGCTTGGGCG	65361
Q	2134	GGCCCGGACATCGAGAGAGCGGGCCAGCCCTG--CCAGGCCCGGGGCAATGAGAGCTTAG	2192
D	65362	GGCCCGGACATCGAGAGAGCGGGCCAGCCCTGAGCCAGGCGCCGGGCAATGAGAGCTTAG	65421
Q	2193	CACCCGGGCAAGGCGCTG--CGAGAGGTGTACTGGGTGGCCGAGAGTCCAGCCGCGCG	2251
D	65422	CACCCGGGCAAGGCGCTGAGGAGGTGTACTGGGTGGCCGAGAGTCCAGAGCGCCGG	65481
Q	2252	CGCTGTGCTGCTGATTTCTCACTGGGCTTACAGAGCTTCCCGCGGGGCAAGGCGCTCG	2311
D	65482	CGCTGTGCTGCTGATTTCTCACTGGGCTTACAGAGCTTCCCGCGGGGCAAGGCGCTCG	65541
Q	2312	GACCTGCAAGCCCGCATCTCTGAGCCTTCCATGGGCTCTGTGTGGCCGAGCCTC	2371
D	65542	GACCTGCAAGCCCGCATCTCTGAGCCTTCCATGGGCTCTGTGTGGCCGAGCCTC	65601
Q	2372	CCCGAGAGCAACACCCCTGCTGCACAGCGCCAGTCCCATGCAACGCAAGGCGTGA	2431
D	65602	CCCGAGAGCAACACCCCTGCTGCACAGCGCCAGTCCCATGCAACGCAAGGCGTGA	65661
Q	2432	GAAGTGCGGGGCAAGCGGCAAGGAGTGGCAGGAGTACCCCTGTCAGACCCCTGTGGCGA	2491
D	65662	GAAGTGCGGGGCAAGCGGCAAGGAGTGGCAGGAGTACCCCTGTCAGACCCCTGTGGCGA	65721
Q	2492	ATCCACTGTGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGAAGCTTGGAGAAGCTTTATG	2551
D	65722	ATCCACTGTGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGAAGCTTGGAGAAGCTTTATG	65781
Q	2552	TCTAGACTAGGGAATGTAATATACAAATACAAATGCAACCCGTGTCTAGGCTCAGAGGCTGGA	2611
D	65782	TCTAGACTAGGGAATGTAATATACAAATACAAATGCAACCCGTGTCTAGGCTCAGAGGCTGGA	65841
Q	2612	ATGCAACAATCACACTCTGTATCTAGTACTGTATGAGGGGCTTGGAGAAGCTTTATGT	2671
D	65842	ATGCAACAATCACACTCTGTATCTAGTACTGTATGAGGGGCTTGGAGAAGCTTTATGT	65901
Q	2672	CTAGCTCAGGGAATGTAATATACAAATACGCACTCTGTATCTAGGCTCAGAGGTTGTAA	2731
D	65902	CTAGCTCAGGGAATGTAATATACAAATACGCACTCTGTATCTAGGCTCAGAGGTTGTAA	65961
Q	2732	CACACCAATCAGCAACCCGTGTCTAGGCTCAGGATGTGGAATGCAACCAATGCAAGTCTG	2791
D	65962	CACACCAATCAGCAACCCGTGTCTAGGCTCAGGATGTGGAATGCAACCAATGCAAGTCTG	66021
Q	2792	TATCTGGCTACTTTCATGGGATCCGTTGTGAAGAGACCAACCAAGGCTTGTGTAGC	2851
D	66022	TATCTGGCTACTTTCATGGGATCCGTTGTGAAGAGACCAACCAAGGCTTGTGTAGC	66081
Q	2852	AATAAAGCTTATATCAGCTGGGTGCAAGGTGGGCTGAGTCCGAAAAGAGTCAAGG	2911
D	66082	AATAAAGCTTATATCAGCTGGGTGCAAGGTGGGCTGAGTCCGAAAAGAGTCAAGG	66141
Q	2912	GAGATAAGGGTGGGCGCGTTTATAGGATTTGGGTAGGTAAAGAAATTTACAGTCAAG	2971
D	66142	GAGATAAGGGTGGGCGCGTTTATAGGATTTGGGTAGGTAAAGAAATTTACAGTCAAG	66201
Q	2972	GGGTTTGTCTCTGGCGGGCAGAGAGTGGGGGTCCGAAGGTCTCAGTGGGGTCTTT	3031
D	66202	GGGTTTGTCTCTGGCGGGCAGAGAGTGGGGGTCCGAAGGTCTCAGTGGGGTCTTT	66261
Q	3032	TTGAGCCAGGATGGCCAGGAAAAGGCTTTCACAAAGTAAATGTCAATTAAGGCAAG	3091
D	66262	TTGAGCCAGGATGGCCAGGAAAAGGCTTTCACAAAGTAAATGTCAATTAAGGCAAG	66321
Q	3092	GACCCGCAATTAACACTCTTTTGTGGTGAATGTATCATAGTTAGTTGG--GCAGGCA	3150
D	66322	GACCCGCAATTAACACTCTTTTGTGGTGAATGTATCATAGTTAGTTGGGNCAGGCA	66381
Q	3151	TATTAATCTTTTGTGTATTTCTTCAATTTACTTCAAGCCATCTGGCGTATGTCCAACT	3210
D	66382	TATTAATCTTTTGTGTATTTCTTCAATTTACTTCAAGCCATCTGGCGTATGTCCAACT	66441
Q	3211	TACAGGGAGTGCATAGGCTTGGGCTGGGCGTCAAGGCTTACAGGCTACCTGGTGGGGCC	3270
D	66442	TACAGGGAGTGCATAGGCTTGGGCTGGGCGTCAAGGCTTACAGGCTACCTGGTGGGGCC	66501
Q	3271	TTGGAGAAATTTTGTGTGCGACACTCTGTATCTAGTTAATCTAGTGGGAGCTGGAGAAC	3330
D	66502	TTGGAGAAATTTTGTGTGCGACACTCTGTATCTAGTTAATCTAGTGGGAGCTGGAGAAC	66561
Q	3331	TTTGTGTCTAGCTCAGGAGTTGTAAACGAC--AATCAGCGCC--TGTCAAAACAGACA	3388
D	66562	TTTGTGTCTAGCTCAGGAGTTGTAAACGACAAATCAGCGCCCTGTCAAAACAGACA	66621
Q	3389	CTGGCTCTACCAATCAGCAGAGT--TGGGTGGGGCCAGATAGAGAAATTAAGCAGGCT	3447
D	66622	CTGGCTCTACCAATCAGCAGAGT--TGGGTGGGGCCAGATAGAGAAATTAAGCAGGCT	66681
Q	3448	GCCCGAGCCAGAGTGGCAAGGCAAGCAGTCCCTATCCCAATATGCGAGCTTGTGCT	3507
D	66682	GCCCGAGCCAGAGTGGCAAGGCAAGCAGTCCCTATCCCAATATGCGAGCTTGTGCT	66741
Q	3508	TTTGGCTTTGGCATAAATCTTGTCTACTGCTCGCTTTT--TGGGTCCACACTGCTTTATG	3566
D	66742	TTTGGCTTTGGCATAAATCTTGTCTACTGCTCGCTTTTGGGTCCACACTGCTTTATG	66801
Q	3567	AGCTGTAACTACACCAAGGCTGTGACCTTCACTCTGGAAGCCACTTAAGCACAG	3626
D	66802	AGCTGTAACTACACCAAGGCTGTGACCTTCACTCTGGAAGCCACTTAAGCACAG	66861
Q	3627	CCGACCGGAGAGTGAACAACTCCGGCGCGCTTGGCTTAAAGCTATTAACACTCACCG	3686
D	66862	CCGACCGGAGAGTGAACAACTCCGGCGCGCTTGGCTTAAAGCTATTAACACTCACCG	6692

QY	398	CACCTACGCGGAGGAGCTAAAGTGGAGATACCTGTAGCCTGGAGGTGAACACTGCAGT	4046
Db	67222	CACCTACGCGGAGGCTAAAGTGGAGATACCTGTAGCCTGGAGGTGAACACTGCAGT	67281
QY	4047	GAGCTGATGTATGACACAGACCCTTAGCCTGGGGACAGACTGAGACCTGTTCCCT	4106
Db	67282	GAGCTGATGTATGACACAGACCCTTAGCCTGGGGACAGACTGAGACCTGTTCCCT	67341
QY	4107	CCGCAAAAAAATTGACAAAAGTGTAAAGAGTGCCTGATATGCTAGGCGCAGTGGCT	4166
Db	67342	CCGCAAAAAAATTGACAAAAGTGTAAAGAGTGCCTGATATGCTAGGCGCAGTGGCT	67401
QY	4167	CATGCCGTGTAATCCGACACTTTGGGAGCCGAGGGGGGGGTACCTAAAGTCAGAG	4226
Db	67402	CATGCCGTGTAATCCGACACTTTGGGAGCCGAGGGGGGGGTACCTAAAGTCAGAG	67461
QY	4227	TGTGAGACGAGCTGGCCAAACATGAGAAAGCCATCTTCTTAAATAACAAATTAGC	4286
Db	67462	TGTGAGACGAGCTGGCCAAACATGAGAAAGCCATCTTCTTAAATAACAAATTAGC	67521
QY	4287	CGGCTGTGGGGGAGTGGTGGAGCATGCTGTAAATCCGATCCAGATCAGAGAGCTAGGCA	4346
Db	67522	CGGCTGTGGGGGAGTGGTGGAGCATGCTGTAAATCCGATCCAGATCAGAGAGCTAGGCA	67581
QY	4347	GGAGAATCACTTGAACCCAGAGAGCGCGCGTTGCAGTGAACGAGATCTGCTATTCAC	4406
Db	67582	GGAGAATCACTTGAACCCAGAGAGCGCGCGTTGCAGTGAACGAGATCTGCTATTCAC	67641
QY	4407	TCCACCCACTCTCGAGCTGGGCAACAGAGCCAAACTCTGTCTTAAAAAATTTTTAAAA	4466
Db	67642	TCCACCCACTCTCGAGCTGGGCAACAGAGCCAAACTCTGTCTTAAAAAATTTTTAAAA	67700
QY	4467	GTCGCTGACATATAAGAGGTGTGCATGCAATGATGTGCCAGCAACATGTTAAGAAATG	4526
Db	67701	GTCGCTGACATATAAGAGGTGTGTGCATGCAATGATGTGCCAGCAACATGTTAAGAAATG	67760
QY	4527	GGAAGCTTCGCTTCCATGAGTGTCTGTTAAAAAACCACCTCAAGGCGAGTGCAGTGGCT	4586
Db	67761	GGAAGCTTCGCTTCCATGAGTGTCTGTTAAAAAACCACCTCAAGGCGAGTGCAGTGGCT	67820
QY	4587	CATGCCATATATATCCGACACTTTGGGAGGCGGAGGGGGTGAATCCGTAGAGTACGAG	4646
Db	67821	CATGCCATATATATCCGACACTTTGGGAGGCGGAGGGGGTGAATCCGTAGAGTACGAG	67880
QY	4647	TTCGAGACGAGCTGCACCAACATGTTGAAATCCACCTCTACTAAAAAATACAAATTT	4706
Db	67881	TTCGAGACGAGCTGCACCAACATGTTGAAATCCACCTCTACTAAAAAATACAAATTT	67940
QY	4707	AGATAGACATGATGTGTGCATGCTGTATATCCCACTACTTTGGAGGCTTAGCGAGAAAA	4766
Db	67941	AGATAGACATGATGTGTGCATGCTGTATATCCCACTACTTTGGAGGCTTAGCGAGAAAA	68000
QY	4767	TCACTATGAACACAGGAGGCGGAGAGTTGTAGTGAACCGAATAGTGGCAATTGGACT--CCAG	4825
Db	68001	TCACTATGAACACAGGAGGCGGAGAGTTGTAGTGAACCGAATAGTGGCAATTGGACT--CCAG	68060
QY	4826	CTGTAGACA--ATGAGCGAAATCCATCTCAAAAAAACACACAAAAAACCACCTCTCTACT	4884
Db	68061	CTGTAGACANATGAGCGAAATCCATCTCAAAAAAACACACAAAAAACCACCTCTCTACT	68120
QY	4885	CCGAGGGAGGTGGGTACA--GAGTGGGCCAATCAATGCAAGGTGTGAGCCACAGAGCT	4943
Db	68121	CCGAGGGAGGTGGGTACANAGTGTGGGCCAATCAATGCAAGGTGTGAGCCACAGAGCT	68180
QY	4944	AAGGCGGAGTCT--CAGGACCGGCGACAGATAAACAGTGTGTGAGATCATGTTGAGAT	5001
Db	68181	AAGGCGGAGTCTGNNKAGAGACCGGGGACAGATPAACAGTGTGTGAGATCATGTTGAGAT	68240
QY	5002	CAGACGCTCCCTGCTATTTGTTGACACACAGGGGGGCCCAAGACACAGAGATGGCCCCATC	5061
Db	68241	CAGACGCTCCCTGCTATTTGTTGACACACAGGGGGGCCCAAGACACAGAGATGGCCCCATC	68300

QY	5062	CAGTCACCAATATCCACTCTCTCATCCAGAAATGTCTTTTGGACAGCTGGGTAATT	5121
Db	68301	CAGTCACCAATATCCACTCTCTCATCCAGAAATGTCTTTTGGACAGCTGGGTAATT	68360
QY	5122	AGGACAGAAAGTACAGTCTGGTGTGGTCAAGTACAGTGCCTCCAGCA-----GGCCTT	5177
Db	68361	AGGACAGAAAGTACAGTCTGGTGTGGTCAAGTACAGTGCCTCCAGCAAGNCCNTT	68420
QY	5178	GTGGCCCTGT--AGAAAACTTCAGGCGCTTAGCGGGGCAAGGTGGC--TCAGCCCTGTATTC	5234
Db	68421	GTGGCCCTGTATGAAAAAGCTTCAGGCGCTTAGCGGGGCAAGGTCGCTCATCGCTGTATTC	68480
QY	5235	CCAGCACTTTGGAGGCGCCAGGCGGGTGGATACAGAGTCAAGAAATGTGACATCTCTG	5294
Db	68481	CCAGCACTTTGGAGGCGCCAGGCGGGTGGATACAGAGTCAAGAAATGTGACATCTCTG	68540
QY	5295	GCTTAACAGGAGTGAACCCCGCTCTCTCTATAAATACAAAAATTTGGCGGGCATGTGTGC	5354
Db	68541	GCTTAACAGGAGTGAACCCCGCTCTCTCTATAAATACAAAAATTTGGCGGGCATGTGTGC	68600
QY	5355	GGGCACTGTAGTTCCAGCTACTCGGAGAGCTGAGGACAGAGAAATGGCGTGAACCCGAGA	5414
Db	68601	GGGCACTGTAGTTCCAGCTACTCGGAGAGCTGAGGACAGAGAAATGGCGTGAACCCGAGA	68660
QY	5415	GGCAGAACTTTGCAAGTACAGCCGAGAAATGCCGCACACTGCATCTCCAGCCTGGGCGCAGAGCA	5474
Db	68661	GGCAGAACTTTGCAAGTACAGCCGAGAAATGCCGCACACTGCATCTCCAGCCTGGGCGCAGAGCA	68720
QY	5475	GACATCCATCTGGAAA--AGAAAAAGAAAAGTTACAGTCTGAGCCAGAGGCCAGAGCTTA	5533
Db	68721	GACATCCATCTGGAAAAGAAAAGAAAAGTTACAGTCTGAGCCAGAGGCCAGAGCTTA	68780
QY	5534	ATTCTGTCACTTACATGACCTTGGGCAAGGCAACTCTCTCCCTGGCCAGTTCAAGGGG	5593
Db	68781	ATTCTGTCACTTACATGACCTTGGGCAAGGCAAGGCAACTCTCTCCCTGGCCAGTTCAAGGGG	68840
QY	5594	TTGGAATCGACTCCAGAGTCCCTCTCCAGCATTTAGCGTGCATGTGTTCT--AAGATGACAG	5652
Db	68841	TTGGAATCGACTCCAGAGTCCCTCTCCAGCATTTAGCGTGCATGTGTTCTNAAGATGACAG	68900
QY	5653	ATGGGGAGATTCCCTCCCTCTCTACGCCAGCCGCTGTCACCTCAAGGGTAATGACACAGG	5712
Db	68901	ATGGGGAGATTCCCTCCCTCTCTACGCCAGCCGCTGTCACCTCAAGGGTAATGACACAGG	68960
QY	5713	AAGTCAGTGTCCCAATCCCGCAGTTCCAAAAGCCCTTGGGCAACCTACTGTCAAGGTGCT	5772
Db	68961	AAGTCAGTGTCCCAATCCCGCAGTTCCAAAAGCCCTTGGGCAACCTACTGTCAAGGTGCT	69020
QY	5773	GCACGAGAGTGAAGTCAAGTCAAGTCAACATGCGCTCGAAGGCTTTGCTCTATTCGGGA	5832
Db	69021	GCACGAGAGTGAAGTCAAGTCAAGTCAACATGCGCTCGAAGGCTTTGCTCTATTCGGGA	69080
QY	5833	CAGCAATCCGGTTCCCTCGTGGCTCTACGCGGAATCTAGGGGCTTTAGCGG--AATGATC	5890
Db	69081	CAGCAATCCGGTTCCCTCGTGGCTCTACGCGGAATCTAGGGGCTTTAGCGGAATGATC	69140
QY	5891	ATGGGGGCGGGGGGGTTCTGGGGGAGTTCCACGTATCAACTTTGGACAGCAAGCC	5950
Db	69141	ATGGGGGCGGGGGGGTTCTGGGGGAGTTCCACGTATCAACTTTGGACAGCAAGCC	69200
QY	5951	TGGAATCTTGCATGTGCTATTCGAAGTGTGGGGTGGGCAAGCAAGCAAGCCAAATGT	6010
Db	69201	TGGAATCTTGCATGTGCTATTCGAAGTGTGGGGTGGGCAAGCAAGCAAGCCAAATGT	69260
QY	6011	CTTATATCTCAAGTAAAGGGCTCAGAGAGTCTCCCAAGACAGGAGCTCCGGAGAGTTTGGG	6070
Db	69261	CTTATATCTCAAGTAAAGGGCTCAGAGAGTCTCCCAAGACAGGAGCTCCGGAGAGTTTGGG	69320
QY	6071	GGTAGGAATGGGAGCAACCAAGGCTTTTCTCTCTCTGAATTTGGGGGCTTGGGGG	6130
Db	69321	GGTAGGAATGGGAGCAACCA--GTTTCTTTTCTCTCTGAATTTGGGGGCTTGGGGG	69379
QY	6131	ACAGGCTTGAATTCACCAAGGAGGAGGGGCAAGAGCACTCCCCCAAGTCT--GGCA	6187


```

/chromosome="17"
/clone="RP11-387H17"
1..1523
misc_feature
    /note="assembly_name:Contig17"
    1624..2830
misc_feature
    /note="assembly_name:Contig18"
    2931..4445
misc_feature
    /note="assembly_name:Contig19"
    4546..5643
misc_feature
    /note="assembly_name:Contig20"
    5744..7022
misc_feature
    /note="assembly_name:Contig21"
    7123..8983
misc_feature
    /note="assembly_name:Contig22"
    9084..11092
misc_feature
    /note="assembly_name:Contig23"
    11193..12752
misc_feature
    /note="assembly_name:Contig24"
    12853..14247
misc_feature
    /note="assembly_name:Contig25"
    14348..16451
misc_feature
    /note="assembly_name:Contig26"
    16552..18402
misc_feature
    /note="assembly_name:Contig27"
    18503..21089
misc_feature
    /note="assembly_name:Contig28"
    21190..24337
misc_feature
    /note="assembly_name:Contig29"
    24438..27881
misc_feature
    /note="assembly_name:Contig30"
    27982..30256
misc_feature
    /note="assembly_name:Contig31
vector_end:sp6
clone_end:slide:left"
30357..32983
misc_feature
    /note="assembly_name:Contig32"
    33084..35459
misc_feature
    /note="assembly_name:Contig33"
    35560..39128
misc_feature
    /note="assembly_name:Contig34"
    39229..42608
misc_feature
    /note="assembly_name:Contig35"
    42709..447067
misc_feature
    /note="assembly_name:Contig36"
    47168..50792
misc_feature
    /note="assembly_name:Contig37"
    50893..55282
misc_feature
    /note="assembly_name:Contig38"
    55383..59386
misc_feature
    /note="assembly_name:Contig39"
    59487..64208
misc_feature
    /note="assembly_name:Contig40"
    64309..68573
misc_feature
    /note="assembly_name:Contig41"
    69074..73514
misc_feature
    /note="assembly_name:Contig42"
    73615..79235
misc_feature
    /note="assembly_name:Contig43"
    79336..84273
misc_feature
    /note="assembly_name:Contig44"
    84374..89567
misc_feature
    /note="assembly_name:Contig45"
    89668..95452
misc_feature
    /note="assembly_name:Contig46"

```

	Query Match	70.1%	Score 4368.4;	DB 2;	Length 22184;
	Best Local Similarity	88.1%;	Pred. No. 0;		
	Matches 4878;	Conservative	0;	Mismatches 621;	Indels 37; Gaps 16;
OY	1	GATCAGTTGGAGCAGTAGTTCGAACACCCTGGGCAGCATGAGGAGACTGTCTCTACG	60		
Ddb	68738	GATCACTTGAGGACAGTAGTTCGAACACCCTGGGCAGCATGAGGAGACTGTCTCTACG	68679		

[illegible]

|||||
Db 67603 CCATTCCTAATTCAGTGTACATATGATACATACTAGGCTCGATCTCTCTCGAG 67544
QY 1197 GCATACCTGAGAGTAGATGTTGCTGTCCAGAGACATACATTTCCACATTAACTAGA 1256
|||||
Db 67543 GCATACCTGAGAGTAGATGTTGCTGTCCAGAGACATACATTTCCACATTAACTAGA 67484
QY 1257 CACTACCAAGTTCATCCCAAGAGAGGTTTTTTTTTACAATCTACACTCCCCAGAAC 1316
|||||
Db 67483 CACTACCAAGTTCATCCCAAGAGAGGTTTTTTTTTACAATCTACACTCCCCAGAAC 67424
QY 1317 AATGAGAGTACTCCAGATCCCTTTACAAGATGCTCTAAGCCCAAGTACAGATGAAAAC 1376
|||||
Db 67423 AATGAGAGTACTCCAGATCCCTTTACAAGATGCTCTAAGCCCAAGTACAGATGAAAAC 67364
QY 1377 AGGAATGGAGGGAGAGTGCAGAGCCCTTTCTAACCATGAGAAATACCTGGTAGACC 1436
|||||
Db 67363 AGGAATGGAGGGAGAGTGCAGAGCCCTTTCTAACCATGAGAAATACCTGGTAGACC 67304
QY 1437 TTCTGATGCTGGAAGAGATGAATTAACGGGGGCTCTGAGAGCTGCCCTGTGAGATCAC 1496
|||||
Db 67303 TTCTGATGCTGGAAGAGATGAATTAACGGGGGCTCTGAGAGCTGCCCTGTGAGATCAC 67244
QY 1497 TGTGACTTCTGAGACCTCCAGTCCAGTCTAGGCCCATGTCATGGCCAGATGAATGAG 1556
|||||
Db 67243 TGTGACTTCTGAGACCTCCAGTCCAG-CTAGCCCATGTGTCAATGGCCAGTGAATGAG 67185
QY 1557 CCCTCACTCTCTGTTGGTCTTTATCT-CCCATGTGGGCTGAAATCTGATGAGCC 1615
|||||
Db 67184 CCCTCACTCTCTGTTGGTCTTTATCTCCCATGTGGGGCTGAAGTGTGAGTGAACC 67125
QY 1616 GTTATTCAGATGTACAGCTTTCTTGACAGAAAGTGTCAACAGAACAGAGGGGCT 1675
|||||
Db 67124 GTTATTCAGATGTACAGCTTTCTTGACAGAAAGTGTCAACAGAACAGAGGGGCT 67065
QY 1676 TGSCAAGATGATTAATCTGCAAAATCTACTGGCTAGCCACAGGATGTGTATCT 1735
|||||
Db 67064 TGSCAAGATGATTAATCTGCAAAATCTACTGGCTAGCCACAGGATGTGTATCT 67005
QY 1736 TGAACAAGTTTTTTCACTTCTCTGAGGCCATCCCTTGCTACAAACACACAGTGTGTTGA 1795
|||||
Db 67004 TGAACAAGTTTTTTCACTTCTCTGAGGCCATCCCTTGCTACAAACACACAGTGTGTTGA 66945
QY 1796 CAGAGTGAATGACG-AGTCCCTTACACTGTAACTCCAGCACTTTGGAGGCCAAAGC 1854
|||||
Db 66944 CAGAGTGAATGACGAAAGTGCCTTACACTGTAACTCCAGCACTTTGGAGGCCAAAGC 66885
QY 1855 GGGTGAATGGCTGAGCCTGAGAGGTGACAGCATGCCGGACATGCTCACAGCCCTGTTTC 1914
|||||
Db 66884 GGGTGAATGGCTGAGCCTGAGAGGTGACAGCATGCCGGACATGCTCACAGCCCTGTTTC 66825
QY 1915 GCTCTGGGGCCTCTCTGCTGGGCTCCCACTTGGTGGCACTTGGAGAGCCCTTACG 1974
|||||
Db 66824 GCTCTGGGGCCTCTCTGCTGGGCTCCCACTTGGTGGCACTTGGAGAGCCCTTACG 66765
QY 1975 CCACCGCTGACATGTGGAGACCCCTTCTGCGCTGGCCAAAGGCCAGCGGCTCCCTCA 2034
|||||
Db 66764 CCACCGCTGACATGTGGAGACCCCTTCTGCGCTGGCCAAAGGCCAGCGGCTCCCTCA 66705
QY 2035 GCTTGAAGGAGGTGTGAGAGGAGGAGGCTCAACAGAGAACCGGGGCTGGCGACGGGCTT 2094
|||||
Db 66704 GCTTGAAGGAGGTGTGAGAGGAGGAGGCTCAACAGAGAACCGGGGCTGGCGAGCGGCTT 66645
QY 2095 GCGGGCCAGCTGAGTTCGCGGTGGGCTGTGGGCTTGGCGGGCCCGCACTCGAGAGAGCG 2154
|||||
Db 66644 GCGGGCCAGCTGAGTTCGCGGTGGGCTGTGGGCTTGGCGGGCCCGCACTCGAGAGAGCG 66585
QY 2155 GCGCAGCCCTGCGAGGCCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGACGGGCTCGGA 2214
|||||
Db 66584 GCGCAGCCCTGCGAGGCCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGAGCGGCTCGGA 66525
QY 2215 GGGTGTACTGGGTGCCCGACAGTGGCAGCCCGCGGCTGTGTGCTCGATTTCTCA 2274
|||||

Db 66524 GGGTGTACTGGGTGCCCGACAGTGGCAGCCCGCGGCTGTGTGCTCGATTTCTCA 66465
QY 2275 CTGGGGCTTATGACAGCCTTCCCGGGGAGGGCTCGGGACCTGACAGCCCGCATGCTGA 2334
|||||
Db 66464 CTGGGGCTTATGACAGCCTTCCCGGGGAGGGCTCGGGACCTGACAGCCCGCATGCTGA 66405
QY 2335 GCGTCCCTTCATAGGGCTCTGTGGGCCCCGAGCCTTCCGAGACACACCCCTGCT 2394
|||||
Db 66404 GCGTCCCTTCATAGGGCTCTGTGGGCCCCGAGCCTTCCGAGACACACCCCTGCT 66345
QY 2395 CCAAGGGCCCATCCCATCCATGAGCAGCAAGGGCTGGAAGTGGGGCGAGGGCACCGG 2454
|||||
Db 66344 CCAAGGGCCCATCCCATCCATGAGCAGCAAGGGCTGGAAGTGGGGCGAGGGCACCGG 66285
QY 2455 GACTGGAGGACACTACCCCTGACAGCCCTGTGTCGGAATCCACTGGGTGAAGCCAGCTGG 2514
|||||
Db 66284 GACTGGAGGACACTACCCCTGACAGCCCTGTGTCGGAATCCACTGGGTGAAGCCAGCTGG 66225
QY 2515 GCTCTGAGTCTGTGGTGAAGACTTGAAGACCTTTATGTCTAGCTCAGGGATGCTAAATAC 2574
|||||
Db 66224 GCTCTGAGTCTGTGGTGAAGACTTGAAGACCTTTATGTCTAGCTCAGGGATGCTAAATAC 66165
QY 2575 ACCAATCAGACCCCTGTGTAGCTCAGGGCTGTGAATGCACCAATCCACACTGTAT 2634
|||||
Db 66164 ACCAATCAGACCCCTGTGTAGCTCAGGGCTGTGTGAATGCACCAATCCACACTGTAT 66105
QY 2635 CTAGCTACTGTATGGGGCCCTTGAGAACCTTTATGTCTAGCTCAGGGATGTAAATACA 2694
|||||
Db 66104 CTAGCTACTGTATGGGGCCCTTGAGAACCTTTATGTCTAGCTCAGGGATGTAAATACA 66045
QY 2695 CCAATGGGCACTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTGC 2754
|||||
Db 66044 CCAATGGGCACTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTGC 65985
QY 2755 TACTCAGGATGTGTAATGCACCAATCGACAGCTGTATCTGCTACTTACAGGGCAT 2814
|||||
Db 65984 TACTCAGGATGTGTAATGCACCAATCGACAGCTGTATCTGCTACTTACAGGGCAT 65925
QY 2815 CCGTGTGAAGAGACACCAAGGCTTTGTGTAGCAATTAAGCTTCTATCACTGGGT 2874
|||||
Db 65924 CCGTGTGAAGAGACACCAAGGCTTTGTGTAGCAATTAAGCTTCTATCACTGGGT 65865
QY 2875 GCAAGTGGGCTGAGTCCGAAAGAGACTCAGCAGAGGAGATTAAGGTGGGCGCTTTTA 2934
|||||
Db 65864 GCAAGTGGGCTGAGTCCGAAAGAGACTCAGCAGAGGAGATTAAGGTGGGCGCTTTTA 65805
QY 2935 TAGGATTTGGGTAGTGAAGGAAATTTACAGTCAAAAGGGGCTTGTCTCTGGCGGGCAG 2994
|||||
Db 65804 TAGGATTTGGGTAGTGAAGGAAATTTACAGTCAAAAGGGGCTTGTCTCTGGCGGGCAG 65745
QY 2995 GAGTGGGGGCTCGCAAGGTGCTCAGTGGGGGCTTTTGTAGCCAGAGATGAGCCAGAAA 3054
|||||
Db 65744 GAGTGGGGGCTCGCAAGGTGCTCAGTGGGGGCTTTTGTAGCCAGAGATGAGCCAGAAA 65685
QY 3055 AGGACTTTCAGAAAGTATGTCAATTAAGGCAAGAGCCGCCATTTACACTCTTTT 3114
|||||
Db 65684 AGGACTTTCAGAAAGTATGTCAATTAAGGCAAGAGCCGCCATTTACACTCTTTT 65625
QY 3115 GTGGTGAATGTATAGTAAAGTTAGTTGGGCGAGGGCAATTCATCTTTTGTGATCTTC 3174
|||||
Db 65624 GTGGTGAATGTATAGTAAAGTTAGTTGGGCGAGGGCAATTCATCTTTTGTGATCTTC 65565
QY 3175 AGTTACTTTCAGGCACTGTGGGCTTATATGTCAAGTTTACAGGGGATGSCATGGCTTGGCT 3234
|||||
Db 65564 AGTTACTTTCAGGCACTGTGGGCTTATATGTCAAGTTTACAGGGGATGSCATGGCTTGGCT 65505
QY 3235 TGGGCTCAGAGGCTTGAACAGCTACTGTGTGGGGCTTGGAGAAATGTTGTGTGACACT 3294
|||||
Db 65504 TGGGCTCAGAGGCTTGAACAGCTACTGTGTGGGGCTTGGAGAAATGTTGTGTGACACT 65445
QY 3295 CTGTATCTAGTTAATCTAGTGGGAGCTGGAAGAACTTTGTGTAGCTCAGGGATTTGA 3354
|||||
Db 65444 CTGTATCTAGTTAATCTAGTGGGAGCTGGAAGAACTTTGTGTAGCTCAGGGATTTGA 65385

QY 3355 AACGCAATTCAGCGCCCTGTCAAAAACAGACCACTGGCTTACCAATCAGCAGATGT 3414
DB 65384 AACGCAATTCAGCGCCCTGTCAAAAACAGACCACTGGCTTACCAATCAGCAGATGT 65325
QY 3415 GGGTGGGGCAGATTAAGAGATTAAGAGCAGGCTGCCGAGCAGCAGTGGCAACGCGCAC 3474
DB 65324 GGGTGGGGCAGATTAAGAGATTAAGAGCAGGCTGCCGAGCAGCAGTGGCAACGCGCAC 65265
QY 3475 AGGTCCCTATCCACAAATATGAGCAGCTTTGTTCTTTGCTGTTGCGATAATCTTGCTAC 3534
DB 65264 AGGTCCCTATCCACAAATATGAGCAGCTTTGTTCTTTGCTGTTGCGATAATCTTGCTAC 65205
QY 3535 TGGTGGCTTTTGGGTCCACACAGCTTTTATGAGCTGTACACTCCACAGAGAGTGTGC 3534
DB 65204 TGGTGGCTTTTGGGTCCACACAGCTTTTATGAGCTGTACACTCCACAGAGAGTGTGC 65145
QY 3595 AGCTTCACTCTCAGAGCAGCTTAAGACACGAGCCACCGGAGGAAATGAACAACCTCCGCG 3654
DB 65144 AGCTTCACTCTCAGAGCAGCTTAAGACACGAGCCACCGGAGGAAATGAACAACCTCCGCG 65085
QY 3655 CGGCTGCTTTAAGAGCTATTAACACTCAGCGGAGGTCTGACGTTCTACTCTCAGCCA 3714
DB 65084 CGGCTGCTTTAAGAGCTATTAACACTCAGCGGAGGTCTGACGTTCTACTCTCAGCCA 65025
QY 3715 GCGAGACCAAGAACCCACAGAAAGAAAGTGCAGAACATCTGAACATCAGAAAGAA 3774
DB 65024 GCGAGACCAAGAACCCACAGAAAGAAAGTGCAGAACATCTGAACATCAGAAAGAA 64965
QY 3775 CAATCTCAGATGCACACCTTAAGAGCTTAACACTCAGTGGAGGGTCCGGGCTTCC 3834
DB 64964 CAATCTCAGATGCACACCTTAAGAGCTTAACACTCAGTGGAGGGTCCGGGCTTCC 64905
QY 3835 TTCTTGAAGTCACTGAGACCAAGCACTCAGCAAGTTTGGACACAGCCAGAGTTTGGAG 3894
DB 64904 TTCTTGAAGTCACTGAGACCAAGCACTCAGCAAGTTTGGACACAGCCAGAGTTTGGAG 64845
QY 3895 ATCAGCTGGGCAACATGATGAATGCCCTCTCGCAAAAAAATTTCAAAAAT 3954
DB 64844 ATCAGCTGGGCAACATGATGAATGCCCTCTCTCTGCG-AAAAAATAAATAATTTCAAAAAT 64786
QY 3955 TGGCGAGCATGGTGGTCCGTGCTGTGTGCCAGTACGCGGAGGCTAAAGTGGAGG 4014
DB 64785 TGGCGAGCATGGTGGTCCGTGCTGTGTGCCAGTACGCGGAGGCTAAAGTGGAGG 64726
QY 4015 ATGCGTTGAGCTTGGGAGGTGAAGACAGTGAAGTGTATTTACCAAGCCCTTAG 4074
DB 64725 ATGCGTTGAGCTTGGGAGGTGAAGACAGTGAAGTGTATTTACCAAGCCCTTAG 64666
QY 4075 GCTGGGGAGACAGTGAAGACCTGTGTTCCCTCGCAAAAAAATTTGACAAAAGTGAATA 4134
DB 64665 GCTGGGGAGACAGTGAAGACCTGTGTTCCCTCGCAAAAAAATTTGACAAAAGTGAATA 64606
QY 4135 AGAGGTGCTGATATGAGTGGCGCAGTGGCTCATGCTGTAATCCAGCAGCTTTGGAA 4194
DB 64605 AGAGGTGCTGATATGAGTGGCGCAGTGGCTCATGCTGTAATCCAGCAGCTTTGGAA 64546
QY 4195 GCGGAGGCGGGGCTCACTAAGGTGAGAGTGTGAGACAGCTTGGCAACATGGAGA 4254
DB 64545 GCGGAGGCGGGGCTCACTAAGGTGAGAGTGTGAGACAGCTTGGCAACATGGAGA 64486
QY 4255 AACCCATCTCTTCTAATAAATAAATAATAGCGGCTGTGGGGGAGTGGGAGCTATG 4314
DB 64485 AACCCATCTCTTCTAATAAATAAATAATAGCGGCTGTGGGGGAGTGGGAGCTATG 64426
QY 4315 CTGTAAATCCAGTACTCAGAGGCTGAGGACAGAGATCATCTTGAACCCAGAGGCGGC 4374
DB 64425 CTGTAAATCCAGTACTCAGAGGCTGAGGACAGAGATCATCTTGAACCCAGAGGCGGC 64366
QY 4375 GGTTCAGTGAAGCCAGATGTGCGCATTTGCACTCAGCCACTCAGCCTGGGCAACAGA 4434
DB 64365 GGTTCAGTGAAGCCAGATGTGCGCATTTGCACTCAGCCACTCAGCCTGGGCAACANN 64306

QY 4435 GCCAACTCTGTCTTAATAAATAAATAAAGTGCCTGCATATAAGAGTGTGCAATG 4494
DB 64305 NNN 64246
QY 4495 CAATAGTGTGAGCAGCAATGTTAAGAAATGTGAGAGTCTGCTT-CCATGCTCTGT 4553
DB 64245 NNN 64186
QY 4554 AAAAACCCACCTCAGAGCAGGTCAGTGGCTCATGCTTAATCCAGCAGCTTTGGGA 4613
DB 64185 TCATTCATTCATTAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 64126
QY 4614 GCGGAGAGCG--GGTGAATACCGCAGCTCAGAGTTCGAGACAGCTGACCAACAGCA 4671
DB 64125 GCGTGAAGTGAAGGAGTATCTTGGCTCACTGCAACCTCTGCTCTGCTGCTCAAGCG 64066
QY 4672 TGGTGAATCCACCTCTACTTAATAAATAAATAATGATGAGATGTTGCTGATGCTG 4731
DB 64065 ATTCTTGGCTCAGCCTCTCTGAGTACTGGGATTAACAGCGGGTGCACACTATGCCAGC 64006
QY 4732 TAATCCACCTACTTGGAGGCTGAGGCAAGAAATCATAGAACAGGAG--GCG 4786
DB 64005 TAATGTTTGTATTTTATAGAGACAGGATTCACCATGTGACGCGCTGCTCGAA 63946
QY 4787 GAGGTGTAGTGAAGCCGAGATCGTCCATTCGACCTCAGCTGAGCAATGAGGAAATC 4846
DB 63945 CTGCTGCTCAAAATTAATCTGCCCCCTTGGCTCCCAAGTGTGAGAGTGGGATTA 63886
QY 4847 CATCTCAAAAACACACACAAAACCCACTCTCTACTCCAGAGGAGCTGGTACAGAGC 4906
DB 63885 CAGGATGAGGACACACACCCGCGCATTCATTCATTAATTTGAAAGTAAAGGCGCTGT 63826
QY 4907 TGGG-----CCATCATGTCAGAGTGTGTCGCCACAGAGCTAAAGCGGAGCTG 4955
DB 63825 TAGGTCTGGGAGACCAAAATATTATTAACTGTGCTCCCACTGTAATAATCTCTAGT 63766
QY 4956 CAGGACGCGGAGACAGATTAACAGTGTGATGATCAGTGTGATGATCAGAGCTGCCGCC 5015
DB 63765 CTATATGGAGATTAATTAACATTAATTTGATTAATTAACAAACATGCTTAATTAACAG 63706
QY 5016 ATGCTGACACACAGGCGGCGCCCAACACAGATGAGCCATCCAGTACACATCC 5075
DB 63705 AACTACTTATGCTGTGATTAAGAAAGCTTTTCATTCCTTCTTACGTGACCTGA 63646
QY 5076 ACTTCTATCAGAGATGTCTGTTCTTGGCAGCGCTGGGTAATTTAGCAGAGGTGA 5135
DB 63645 GTTGTGATTTGTGACATCAGAAATGTTTATGCAATTTGATTTGATGACAGTTTGG 63586
QY 5136 CAGTCTTGGGAGTGTGATCAGACTGCCAGGACGCTTGGGCTGTAGAAAAGT 5195
DB 63585 AATTTTACTAAGATTTATTTTAACATGCTTTTAACCTTTTCTTCTGTTTTTAA 63526
QY 5196 TCAAGCCTAGCGGCGGACAGGCTCAGCTCCTGTAAATCCAGCAGCTTTGGAGCGCAG 5255
DB 63525 AGATTAAGTGTGCGGCGGCTGACCGCTGTAATCTTACACTTTGGAGGCGCAG 63466
QY 5256 GCGGCTGATCA--CGAGTTCAGAGATGTGTACACTCTGCTTAACAGCTGAACCCC 5313
DB 63465 GTGGGCGATCATCTGAGGTCCAGAGTTCCAGACACACCTGACCAACATGAGAAACCT 63406
QY 5314 GTCTCTACTAATAAATAAATAAATGCGGCGGATGTTGGCGGACCTGTAGTTCAGC 5373
DB 63405 GCTCTACTAATAAATGCG--AAATTTAGCAGGCGGATGTTGATCATGCTTAATCCAGC 63348
QY 5374 TACTCGGAGGCTGAGGACAGAGATGCTGAACCCGAGAGGCGAGATTTGAGTGAAGC 5433
DB 63347 TACTCGGAGGCTGAGGACAGAGATGCTGAACCCGAGAGGCGAGTGTGCTGAGC 63288
QY 5434 CGAGATGCGCCACTGCACTCAGCTGAGCGAGC-AGAGCAAGACTCATCTGGAAGAA 5492
DB 63287 CGAGATGCGCCACTGCACTCAGCTGAGCGAGCGAGAGGAGAACTCTGCTCAAAAAA 63228
QY 5493 AAAAGAAACGTTAG 5508

Db 63227 AAAAAAAAAAGATGAG 63212

RESULT 9
LOCUS AR156465 2834 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 6 from patent US 6242218.
ACCESSION AR156465
VERSION AR156465.1 GI:15125169
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2834)
AUTHORS Treco,D.A., Heartlein,M.W. and Selden,R.F.
TITLE Genomic sequences for protein production and delivery
JOURNAL Patent: US 6242218-A 6 05-JUN-2001;
FEATURES
source location/Qualifiers
1..2834
BASE COUNT 670 a /organism="unknown" 588 t
ORIGIN
Query Match 45.5%; Score 2834; DB 6; Length 2834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1890 CCGGAGTCCCTCAGACCCCTCGTTCGCTCGCGGCTCCTCGCTGGGCTCCACTTC 1949
Db 1 CCGGAGTCCCTCAGACCCCTCGTTCGCTCGCGGCTCCTCGCTGGGCTCCACTTC 60

QY 1950 GGTGGCACTTGAAGAGCCCTTCAGCCACCCTGCACTGTGGAGCCCTTCTGGGCTG 2009
Db 61 GGTGGCACTTGAAGAGCCCTTCAGCCACCCTGCACTGTGGAGCCCTTCTGGGCTG 120

QY 2010 GCCAAGCCAGAGCCGGCTCCCTCAGCTTGGAGGAGGTGTGAGGAGAGAGGCTCAAGCA 2069
Db 121 GCCAAGCCAGAGCCGGCTCCCTCAGCTTGGAGGAGGTGTGAGGAGAGGCTCAAGCA 180

QY 2070 GGAACCGGGGCTGCGCAGCGGCTTGCGGCCAGCTGAGTTCCGGGTGGGCTTG 2129
Db 181 GGAACCGGGGCTGCGCAGCGGCTTGCGGCCAGCTGAGTTCCGGGTGGGCTTG 240

QY 2130 GGGGGGCCCCGCACTGGAGCAGCGGGCCAGCCCTGCCAGGCCCGGCAATGAGAGCT 2189
Db 241 GGGGGGCCCCGCACTGGAGCAGCGGGCCAGCCCTGCCAGGCCCGGCAATGAGAGCT 300

QY 2190 TAGACCCGGGGCAGGGGCTGGGAGGGTGTACTGGGTCCCGACAGATGCCAGCCGCC 2249
Db 301 TAGACCCGGGGCAGGGGCTGGGAGGGTGTACTGGGTCCCGACAGATGCCAGCCGCC 360

QY 2250 GGGGCTGTGCTGCTGATTTCTCACTGGGCTTACAGACCTTCCGCGGGGAGGGCTC 2309
Db 361 GGGGCTGTGCTGCTGATTTCTCACTGGGCTTACAGACCTTCCGCGGGGAGGGCTC 420

QY 2310 GGGAGCTGCAAGCCCGCATGCTGAGGCTTCCCTTCATGGGCTCTGTGCGGCGCAGCC 2369
Db 421 GGGAGCTGCAAGCCCGCATGCTGAGGCTTCCCTTCATGGGCTCTGTGCGGCGCAGCC 480

QY 2370 TCCCCAGACAGACACCCCTGCTCAAGGCGCCAGTCCCATGACACAGAGGGCT 2429
Db 481 TCCCCAGACAGACACCCCTGCTCAAGGCGCCAGTCCCATGACACAGAGGGCT 540

QY 2430 GAGAAGTGGGGGAGCGGACCGGGAGTGGGAGGAGTACCCCTGCAAGCCCTGGTGG 2489
Db 541 GAGAAGTGGGGGAGCGGACCGGGAGTGGGAGGAGTACCCCTGCAAGCCCTGGTGG 600

QY 2490 GAATCCACTGGGTGAAGCAGCTGGGCTCTGAGTCTGTGAGACTTGAAGAACTTTA 2549
Db 601 GAATCCACTGGGTGAAGCAGCTGGGCTCTGAGTCTGTGAGACTTGAAGAACTTTA 660

QY 2550 TGTCTAGCTCAGGAGTGTAAATACACAATGACAGACCTGTGTACTGAGGTCTGT 2609

Db 661 TGTCTAGCTCAGGAGTGTAAATACACAATGACAGACCTGTGTACTGAGGTCTGT 720

QY 2610 GAATGACCAATCCACACTCTGTATCTAGTACTGATGGGCTTGGAGAACTTTAT 2669
Db 721 GAATGACCAATCCACACTCTGTATCTAGTACTGATGGGCTTGGAGAACTTTAT 780

QY 2670 GTCTAGCTCAGGAGTGTAAATACACAATCCGACCTGTATCTAGTCAAGTTGTA 2729
Db 781 GTCTAGCTCAGGAGTGTAAATACACAATCCGACCTGTATCTAGTCAAGTTGTA 840

QY 2730 AACACCAATCAGCAACCCCTGTCTAGCTCAGGAGTATGTGAATGACCAATCGAAGTC 2789
Db 841 AACACCAATCAGCAACCCCTGTCTAGCTCAGGAGTATGTGAATGACCAATCGAAGTC 900

QY 2790 TGTATCTGGCTACTTTCATATGAGGCTCGTGTGAAGAGACACCAACAGGCTTGTGTA 2849
Db 901 TGTATCTGGCTACTTTCATATGAGGCTCGTGTGAAGAGACACCAACAGGCTTGTGTA 960

QY 2850 GCATTAAGCTTCTATCACCCTGGGTGACAGTGGGCTGCCAAAAAGAGTCAAGCAA 2909
Db 961 GCATTAAGCTTCTATCACCCTGGGTGACAGTGGGCTGCCAAAAAGAGTCAAGCAA 1020

QY 2910 GGGAGATAAGGGTGGGCGGCTTTATAGGATTTGGGTAGTTAAGGAAATTAAGTCA 2969
Db 1021 GGGAGATAAGGGTGGGCGGCTTTATAGGATTTGGGTAGTTAAGGAAATTAAGTCA 1080

QY 2970 AGGGGGTTTGTCTCTGGGGGAGAGTGGGGGCTGCGCAAGTGTCTCAGTGGGGTGTCT 3029
Db 1081 AGGGGGTTTGTCTCTGGGGGAGAGTGGGGGCTGCGCAAGTGTCTCAGTGGGGTGTCT 1140

QY 3030 TTTTGAAGCAGATGAGCCAGGAAGAAGACTTTCACAGGTAATGTCAATTAAGCA 3089
Db 1141 TTTTGAAGCAGATGAGCCAGGAAGAAGACTTTCACAGGTAATGTCAATTAAGCA 1200

QY 3090 AGGACCGGCATTTAACCTTTTGTGGTGAATGTCACTTAAGTTGGGGCAGGGC 3149
Db 1201 AGGACCGGCATTTAACCTTTTGTGGTGAATGTCACTTAAGTTGGGGCAGGGC 1260

QY 3150 ATATTACATCTTTTGTGATTTCTCAGTACTTCAAGGCATCTGGGCGTATATGTGCAAG 3209
Db 1261 ATATTACATCTTTTGTGATTTCTCAGTACTTCAAGGCATCTGGGCGTATATGTGCAAG 1320

QY 3210 TTACAGGGAGTGGATGCTTGGCTTGAGGCTCAGAGGCTTGAAGCTTCTGTGGGCG 3269
Db 1321 TTACAGGGAGTGGATGCTTGGCTTGAGGCTCAGAGGCTTGAAGCTTCTGTGGGCG 1380

QY 3270 CTTTGAAGATCTTTGTGTGACACCTGTATCTAGTTATCTAGTGGGAGCTGGAGAAC 3329
Db 1381 CTTTGAAGATCTTTGTGTGACACCTGTATCTAGTTATCTAGTGGGAGCTGGAGAAC 1440

QY 3330 CTTTGTGTAGCTCAGGAGTGTAAACGACCAATCAGCGCCCTGTCAAAACAGACGAC 3389
Db 1441 CTTTGTGTAGCTCAGGAGTGTAAACGACCAATCAGCGCCCTGTCAAAACAGACGAC 1500

QY 3390 TCGGCTCTACCAATCAGAGAGATGTGGGTGGGCGCAGATTAAGATAAAGCAGGCTGC 3449
Db 1501 TCGGCTCTACCAATCAGAGAGATGTGGGTGGGCGCAGATTAAGATAAAGCAGGCTGC 1560

QY 3450 CCGAGCCAGCAGTGGCAGACGCGACAGGTCCCTATCCACAAATATGGCAGCTTGTCTTT 3509
Db 1561 CCGAGCCAGCAGTGGCAGACGCGACAGGTCCCTATCCACAAATATGGCAGCTTGTCTTT 1620

QY 3510 TGGTGTGGCGAATTAATCTTGTACTGCTGCGCTTTTGGGTGCCACAGTCTTTTATGAGC 3569
Db 1621 TGGTGTGGCGAATTAATCTTGTACTGCTGCGCTTTTGGGTGCCACAGTCTTTTATGAGC 1680

QY 3570 TGTAAACATCACCAGGAAGTCTGCAAGCTTCACTCCTGAAGGCAATGAGACCAAGGCC 3629
Db 1681 TGTAAACATCACCAGGAAGTCTGCAAGCTTCACTCCTGAAGGCAATGAGACCAAGGCC 1740

QY 3630 ACCGGGAGGAATGAACAATCGGCGGCTGCTTAAGAGCTATTAAGCTCACCAGCA 3689
Db 1741 ACCGGGAGGAATGAACAATCGGCGGCTGCTTAAGAGCTATTAAGCTCACCAGCA 1800

Dh 1741 ACCGGAGGAGATGAACAACCTCCGCCGCTGCTTAAGAGCTATTAACCTACCGCGAA 1800
Qy 3690 GGTCTGCAGCTTCACTCTCACCAGCGAGACCAAGAACCCACAGAGAGAACTGC 3749
Db 1801 GGTCTGCAGCTTCACTCTCACCAGCGAGACCAAGAACCCACAGAGAGAACTGC 1860
Qy 3750 GAACACATCTGAACATCAGAGAGAACAACTCCAGATGACCACTTAAGAGCTGTAACA 3809
Db 1861 GAACACATCTGAACATCAGAGAGAACAACTCCAGATGACCACTTAAGAGCTGTAACA 1920
Qy 3810 CTCACCTGCGAGGTCGCGGCTTCTTCTTGAAGTCACTAGAGACCAAGCACTCACCAGTT 3869
Db 1921 CTCACCTGCGAGGTCGCGGCTTCTTCTTGAAGTCACTAGAGACCAAGCACTCACCAGTT 1980
Qy 3870 TCGGACACAAAGCCAGAGAGTTGAGATCAGCTGGGCAACATGATGAATGCCCTCTCG 3929
Db 1981 TCGGACACAAAGCCAGAGAGTTGAGATCAGCTGGGCAACATGATGAATGCCCTCTCG 2040
Qy 3930 CAAAAAATTTACAAAAATTTGGCGAGACATGATGTCCTGCTGTCCTGTCCTG 3989
Db 2041 CAAAAAATTTACAAAAATTTGGCGAGACATGATGTCCTGCTGTCCTGTCCTG 2100
Qy 3990 CTACGCGGAGGCTAAAGTGGAGAGATCGCTTGAAGCTGGAGAGTGAAGACTGCACTGAG 4049
Db 2101 CTACGCGGAGGCTAAAGTGGAGAGATCGCTTGAAGCTGGAGAGTGAAGACTGCACTGAG 2160
Qy 4050 CTGATATTTACCACAGCCCTTAGGCTGGGGACAGACTGAGACCTCTTCCCTCCG 4109
Db 2161 CTGATATTTACCACAGCCCTTAGGCTGGGGACAGACTGAGACCTCTTCCCTCCG 2220
Qy 4110 CAAAAAATTTGACAAAAATTTGAAGAGAGTCTGATATGCTAGGCGCAGTGCCTCAT 4169
Db 2221 CAAAAAATTTGACAAAAATTTGAAGAGAGTCTGATATGCTAGGCGCAGTGCCTCAT 2280
Qy 4170 GCTCTGATCCAGCACTTTGGGAAGCCGAGCGGGGCTCACTAAGTCAAGAGTGT 4229
Db 2281 GCTCTGATCCAGCACTTTGGGAAGCCGAGCGGGGCTCACTAAGTCAAGAGTGT 2340
Qy 4230 GAGACAGGCTGGCCAACTGGAGAAAGCCCATCTTTAAATAATCAAAATTAAGCCG 4289
Db 2341 GAGACAGGCTGGCCAACTGGAGAAAGCCCATCTTTAAATAATCAAAATTAAGCCG 2400
Qy 4290 CTGTGGGGGCACTGTGTGAGCATGCTGTAATCCAGTACTCAGAGGCTGAGGAGAGA 4349
Db 2401 CTGTGGGGGCACTGTGTGAGCATGCTGTAATCCAGTACTCAGAGGCTGAGGAGAGA 2460
Qy 4350 GAATCACTTGAACCCAGAGGCGGCGGTTGCAGTGAGCCGAGATCGTGCATTGCACTCC 4409
Db 2461 GAATCACTTGAACCCAGAGGCGGCGGTTGCAGTGAGCCGAGATCGTGCATTGCACTCC 2520
Qy 4410 ACCCACTCCAGGCTGGGCAACAAGAGCCAACTCTCTTAAATAATTAATTAATTAATTA 4469
Db 2521 ACCCACTCCAGGCTGGGCAACAAGAGCCAACTCTCTTAAATAATTAATTAATTAATTA 2580
Qy 4470 CCTGACATTAAGAGAGTGTGCAATGATGTCAGAGCAACATGTTAAAGATTTGA 4529
Db 2581 CCTGACATTAAGAGAGTGTGCAATGATGTCAGAGCAACATGTTAAAGATTTGA 2640
Qy 4530 GCTCTGCTTCCATGCTCTCTTAAATAATTAATTAATTAATTAATTAATTAATTAATTA 4589
Db 2641 GCTCTGCTTCCATGCTCTCTTAAATAATTAATTAATTAATTAATTAATTAATTAATTA 2700
Qy 4590 GCCTAATATCCAGACATTTGGAGGCGAGGCGGTTGATATCACTGAGAGTCAAGAGTTC 4649
Db 2701 GCCTAATATCCAGACATTTGGAGGCGAGGCGGTTGATATCACTGAGAGTCAAGAGTTC 2760
Qy 4650 GAGACAGGCTGACCAACCAATGATGTAATCCACCTCTACTATAAATAATTAATTAATTA 4709
Db 2761 GAGACAGGCTGACCAACCAATGATGTAATCCACCTCTACTATAAATAATTAATTAATTA 2820
Qy 4710 TGAGCATGCTGCTG 4723
Db 2821 TGAGCATGCTGCTG 2834

RESULT 10
AF388025 5527 bp DNA linear PRI 14-JUN-2001
LOCUS Homo sapiens colony stimulating factor 3 (granulocyte) (CSF3) gene,
DEFINITION complete cds.
ACCESSION AF388025
VERSION AF388025.1 GI:14423572
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5527)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattleSNPs, NHLBI Program for
Genomic Applications, UW-FHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu).
FEATURES
source location/Qualifiers
1..5527
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
14..315
/rpt_family="AluX"
/rpt_type=dispersed
28
/frequency="0.43"
/replace="A"
321
/frequency="0.26"
/replace="T"
379
/frequency="0.04"
/replace="T"
649..944
/rpt_family="AluY"
/rpt_type=dispersed
789
/frequency="0.06"
/replace="T"
956..1089
/rpt_family="MIR3"
/rpt_type=dispersed
1158
/frequency="0.18"
/replace="A"
1577
/frequency="0.01"
/replace="A"
1607
/frequency="0.17"
/replace="C"
1619
/frequency="0.01"
/replace="A"
1980
/frequency="0.49"
/replace="G"
2005..4378
/gene="CSF3"
join(2005..2079,2256..2419,2798..2905,3050..3196,
3361..4378)
/gene="CSF3"
Product="colony stimulating factor 3 (granulocyte)"
2024
/gene="CSF3"
/frequency="0.02"

CDS
/replace="G"
join(2040..2079,2256..2419,2798..2905,3050..3196,
3361..3525)
/gene="CSF3"
/codon_start=1
/product="colony stimulating factor 3 (granulocyte)"
/protein_id="AAK62469.1"
/db_xref="GI:14423573"
/translation="MAGPATQSPKIMALQLLMHSALMTVOEATPLGPASSLPQSFLLKCLEQYRKIQGDGAAIQEKLVSQAYKLCHEIYLGHSLGIPAPLSSCPGQALQLAGLSLSHGLFYQGLALEGISELGPPLTDQLDVADEATTIWQMEELGMAPALOPTGAMPAPASAFQRRAGVIVASHQSFLEVSRYLRHLAQP"
2504
variation
/gene="CSF3"
/frequency="0.07"
/replace="G"
2584
variation
/gene="CSF3"
/frequency="0.08"
/replace="T"
2589
variation
/gene="CSF3"
/frequency="0.01"
/replace="C"
2617
variation
/gene="CSF3"
/frequency="0.05"
/replace="T"
2725
variation
/gene="CSF3"
/frequency="0.01"
/replace="C"
2734
variation
/gene="CSF3"
/frequency="0.01"
/replace="A"
2764
variation
/gene="CSF3"
/frequency="0.03"
/replace="T"
3254
variation
/gene="CSF3"
/frequency="0.03"
/replace="A"
3370
variation
/gene="CSF3"
/frequency="0.03"
/replace="A"
3421
variation
/gene="CSF3"
/frequency="0.02"
/replace="A"
3456
variation
/gene="CSF3"
/frequency="0.16"
/replace="A"
3686
variation
/gene="CSF3"
/frequency="0.03"
/replace="T"
4050
variation
/gene="CSF3"
/frequency="0.23"
/replace="T"
4105
variation
/gene="CSF3"
/frequency="0.03"
/replace="C"
4145
variation
/gene="CSF3"
/frequency="0.01"
/replace="A"
4215

/gene="CSF3"
/frequency="0.47"
/replace="T"
4480
variation
/frequency="0.16"
/replace="G"
4522
variation
/frequency="0.02"
/replace="T"
4530..4762
repeat_region
/rpt_family="L2"
/rpt_type=dispersed
4993
variation
/frequency="0.02"
/replace="T"
5168
variation
/frequency="0.47"
/replace="G"
BASE COUNT 1195 a 1571 c 1673 g 1088 t
ORIGIN

Query Match 25.9%; Score 1612.4; DB 9; Length 5527;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 11; Indels 4; Gaps 4;
QY 4557 AACCCACCTCAAGCCAGGTGCAGTGCTCATCTTAATCCACACTTTGGAGGC 4616
Db 1 AACCCACCTCAAGCCAGGTGCAGTGCTCATCTTAATCCACACTTTGGAGGC 60
QY 4617 CGAGCGGGGATCAGCTGAGTCAGAGTTCGAGACCGCCAGCACCAACATGTTG 4676
Db 61 CGAGCGGGGATCAGCTGAGTCAGAGTTCGAGACCGCCAGCACCAACATGTTG 120
QY 4677 AATCCACCTCTACTAAAAATACAAATTAATGATGACATGTTGTCATGCTGTATTC 4736
Db 121 AATCCACCTCTACTAAAAATACAAATTAATGATGACATGTTGTCATGCTGTATTC 180
QY 4737 CCACCTACTTGGAGGCTGAGGCGAGAAATCACTAAGACGAGGAGCGGAGTTGTAG 4796
Db 181 CCACCTACTTGGAGGCTGAGGCGAGAAATCACTAAGACGAGGAGCGGAGTTGTAG 240
QY 4797 TGAGCCGAGATCGTCCATTGATCACTCCAGCTGAGCAATGAGCAAACTCCATCTCAAA 4856
Db 241 TGAGCCGAGATCGTCCATTGATCACTCCAGCTGAGCAATGAGCAAACTCCATCTCAAA 300
QY 4857 AACACACACAAAAACCACTCTCTACTCCAGGAGCTGGGTACAGAGCTGGGCCAAT 4916
Db 301 AACACACACAAAAACCACTCTCTACTCCAGGAGCTGGGTACAGAGCTGGGCCAAT 360
QY 4917 CAGTGCAGAGTCTGAGCCACAGACTTAAGCGGAGCTGCAGACCGCGGACCGATTAAC 4976
Db 361 CAGTGCAGAGTCTGAGCCACAGACTTAAGCGGAGCTGCAGACCGCGGACCGATTAAC 420
QY 4977 AGTGTGAGATCGATGTTGTTGATCAAGAGCTCCCTGCATTTGTTGACCAACCAAGGGGCC 5036
Db 421 AGTGTGAGATCGATGTTGTTGATCAAGAGCTCCCTGCATTTGTTGACCAACCAAGGGGCC 480
QY 5037 CCCAAGCAGCAGAGATGGCCCATCAGTCCACATCCACTTCTCATCCAGAGATGCT 5096
Db 481 CCCAAGCAGCAGAGATGGCCCATCAGTCCACATCCACTTCTCATCCAGAGATGCT 540
QY 5097 GTTTCCTGGCAGCGTGGGGTAAATTAGACAGAAAGTTCAGGCTTAAGGCCGGGCAAGTC 5156
Db 541 GTTTCCTGGCAGCGTGGGGTAAATTAGACAGAAAGTTCAGGCTTAAGGCCGGGCAAGTC 600
QY 5157 AGATGCCCCAGGAGCGCTTGTGCTGTAGAAAAGTTCAGGCTTAAGGCCGGGCAAGTC 5216
Db 601 AGATGCCCCAGGAGCGCTTGTGCTGTAGAAAAGTTCAGGCTTAAGGCCGGGCAAGTC 660
QY 5217 TGCTCAGCGCTGTAAATCCAGCACTTTGGAGGCGGAGGCGGTTGATCAGGAGTCAG 5276
Db 661 T-GCTCAGCGCTGTAAATCCAGCACTTTGGAGGCGGAGGCGGTTGATCAGGAGTCAG 719


```
* 30257 30356: gap of unknown length
* 30357 30356: contig of 2627 bp in length
* 32984 32983: gap of unknown length
* 33083 33083: contig of 2376 bp in length
* 33084 33083: gap of unknown length
* 35459 35459: gap of unknown length
* 35460 35459: gap of unknown length
* 35559 35559: gap of unknown length
* 39128 39128: contig of 3569 bp in length
* 39129 39228: gap of unknown length
* 39229 39228: gap of unknown length
* 42608 42608: contig of 3380 bp in length
* 42609 42708: gap of unknown length
* 42709 42708: gap of unknown length
* 47068 47067: contig of 4359 bp in length
* 47068 47167: gap of unknown length
* 50793 50792: contig of 3625 bp in length
* 50793 50893: gap of unknown length
* 50893 55282: contig of 4390 bp in length
* 55283 55382: gap of unknown length
* 55383 59386: contig of 4004 bp in length
* 59387 59486: gap of unknown length
* 59487 64208: contig of 4722 bp in length
* 64209 64308: gap of unknown length
* 64309 68973: contig of 4665 bp in length
* 68974 69073: gap of unknown length
* 69074 73514: contig of 4441 bp in length
* 73515 73614: gap of unknown length
* 73615 79235: contig of 5621 bp in length
* 79236 79335: gap of unknown length
* 79336 84273: contig of 4938 bp in length
* 84274 84373: gap of unknown length
* 84374 89567: contig of 5194 bp in length
* 89568 89667: gap of unknown length
* 89668 95452: contig of 5785 bp in length
* 95453 95552: gap of unknown length
* 95553 101367: contig of 5815 bp in length
* 101368 101467: gap of unknown length
* 101468 108294: contig of 6827 bp in length
* 108295 108394: gap of unknown length
* 108395 115531: contig of 7137 bp in length
* 115532 115631: gap of unknown length
* 115632 125051: contig of 9420 bp in length
* 125052 125151: gap of unknown length
* 125152 134705: contig of 9554 bp in length
* 134706 134805: gap of unknown length
* 134806 145490: contig of 10685 bp in length
* 145491 145590: gap of unknown length
* 145591 155823: contig of 10233 bp in length
* 155824 155923: gap of unknown length
* 155924 170624: contig of 14701 bp in length
* 170625 170724: gap of unknown length
* 170725 184829: contig of 14105 bp in length
* 184830 203099: gap of unknown length
* 203099 203199: contig of 18170 bp in length
* 203100 221484: gap of unknown length
* 221484 221484: contig of 18285 bp in length.

FEATURES
    source
        1. .221484
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="17"
            /clone="RP11-387H17"
            1. .1523
                /note="assembly_name:Contig17"
                /note="assembly_name:Contig18"
                /note="assembly_name:Contig19"
                /note="assembly_name:Contig20"
                /note="assembly_name:Contig21"
                /note="assembly_name:Contig22"
                /note="assembly_name:Contig23"
                /note="assembly_name:Contig24"
                /note="assembly_name:Contig25"
                /note="assembly_name:Contig26"
                /note="assembly_name:Contig27"
                /note="assembly_name:Contig28"
                /note="assembly_name:Contig29"
                /note="assembly_name:Contig30"
                /note="assembly_name:Contig31"
                /note="assembly_name:Contig32"
                /note="assembly_name:Contig33"
                /note="assembly_name:Contig34"
                /note="assembly_name:Contig35"
                /note="assembly_name:Contig36"
                /note="assembly_name:Contig37"
                /note="assembly_name:Contig38"
                /note="assembly_name:Contig39"
                /note="assembly_name:Contig40"
                /note="assembly_name:Contig41"
                /note="assembly_name:Contig42"
                /note="assembly_name:Contig43"
                /note="assembly_name:Contig44"
                /note="assembly_name:Contig45"
                /note="assembly_name:Contig46"

Query Match      22.9%; Score 1424.8; DB 2; Length 221484;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1480; Conservative 0; Mismatches 17; Indels 5; Gaps 4;

QY 4732 TAAATCCACCTACTTGGAGAGCTGAGCGAGAAATCAGTGAACCGAGGAGCGAGGT 4791
DB 1 TAAATCCACCTACTTGGAGAGCTGAGCGAGAAATCAGTGAACCGAGGAGCGAGGT 60
QY 4792 TGTAGTGAGCGGAGATGCTCCATTCGACCTTCAGCAGCTGAGCAATGACGAACTCCATCT 4851
DB 61 TGTAGTGAGCGGAGATGCTCCATTCGACCTTCAGCAGCTGAGCAATGACGAACTCCATCT 120
QY 4852 CAAAAAACAACAAACAAACAAACCACTCTCTACCTCCAGGAGAGCTGGGTACAGAGCTGGGC 4911
DB 121 CAAAAAACAACAAACAAACCAACCTCTCTACCTCCAGGAGAGCTGGGTACAGAGCTGGGC 180
QY 4912 CACATCAGTCAAGGTGCTGAGCCACAGAGACTAAGGGGAGCTGCAAGGAGCGGAGCCAG 4971
DB 181 CACATCAGTCAAGGTGCTGAGCCACAGAGACTAAGGGG--ACTGCAAGGAGCGGAGCCAG 238
QY 4972 ATAACAGTGTGTGAGATCACTGTGTGAGATCAGACGCTCCCTGCATTGTGTGACACACAG 5031
DB 239 ATAACAGTGTGTGAGATCACTGTGTGAGATCAGACGCTCCCTGCATTGTGTGACACACAG 298
```


|||||

Db 541 CGAGGTGAGGATGCTGATCCATCTGCTGCTAACGCTGTAACCCCGTCTCTACTATAA 600

QY 5338 TACAAAAAATGGCCGGGCGATGCTGGCGGCGACCTGTAGTTCCAGCTACTCGGGAGCTG 5387

Db 601 TACAAAAAATGGCCGGGCGATGCTGGCGGCGACCTGTAGTTCCAGCTACTCGGGAGCTG 660

QY 5388 AGCGAGGAATGGGCGTGAACCCGAGAGAGAGTTTTCAGAGTCCGAGCCGAGC 5447

Db 661 AGCGAGGAATGGGCGTGAACCCGAGAGAGTTTTCAGAGTCCGAGCCGAGC 720

QY 5448 TGCACCTCCAGCTGGGCGACAGAGCAAGTCTCATCTGGAAGAGAGAGAGAGTTC 5507

Db 721 TGCACCTCCAGCTGGGCGACAGAGCAAGTCTCATCTGGAAGAGAGAGAGTTC 780

QY 5508 GGTCTGAGCCAGAGGCGCCAGGCTGTAACTTGTCTCATTACATGACTTTGGGCAAGC 5567

Db 781 GGTCTGAGCCAGAGGCGCCAGGCTGTAACTTGTCTCATTACATGACTTTGGGCAAGC 840

QY 5568 TTCCCTCCGCTGGCCAGTTACAGGGGGTGGATGACATCCAAAGGCTCTCCAGATTA 5627

Db 841 TTCCCTCCGCTGGCCAGTTACAGGGGGTGGATGACATCCAAAGGCTCTCCAGATTA 900

QY 5628 CGCTGACATGTTCTAAGATGAGAGATGGGCGAGTTCCCTCTCTCAAGCCAGCCGCTG 5687

Db 901 CGCTGACATGTTCTAAGATGAGAGATGGGCGAGTTCCCTCTCTCAAGCCAGCCGCTG 960

QY 5688 TCCACTTCAGGTGAATGAGAGAGAGTCACTGCTCCCAATCCCGAGTTCCAAAGCC 5747

Db 961 TCCACTTCAGGTGAATGAGAGAGAGTCACTGCTCCCAATCCCGAGTTCCAAAGCC 1020

QY 5748 TTGGGAGACCTACTGTGTCAGGCTGTCGACGAGAGAGTGAAGTTCAGGTGAGCC 5807

Db 1021 TTGGGAGACCTACTGTGTCAGGCTGTCGACGAGAGAGTGAAGTTCAGGTGAGCC 1080

QY 5808 TCGAAGGCTTTCCTCATTCCTGCGAGACATCCGCTTCTCTGCTCTACCGGATTC 5867

Db 1081 TCGAAGGCTTTCCTCATTCCTGCGAGACATCCGCTTCTCTGCTCTACCGGATTC 1140

QY 5868 TAGGGGCTTTAGCGAATGATGATGAGGGGGGGGGGGGTTTCGGGGAGTTCACGCT 5927

Db 1141 TAGGGGCTTTAGCGAATGATGATGAGGGGGGGGGGGGTTTCGGGGAGTTCACGCT 1200

QY 5928 AATCACTTGGGAGAGACACCTGTGAATCTTCATGCTGCTATCCAAAGT 5979

Db 1201 AATCACTTGGGAGAGACACCTGTGAATCTTCATGCTGCTATCCAAAGT 1252

RESULT 13

BX470173

LOCUS 110779 bp DNA linear PRI 21-MAY-2003

DEFINITION Human DNA sequence from clone Rp11-739C15 on chromosome 1, complete

ACCESSION BX470173

VERSION BX470173.5 GI:31043711

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 110779)

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE Githero, R.

JOURNAL Direct Submission

COMMENT Submitted (21-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk On May 22, 2003 this sequence version replaced gi:30962389.

Center: Wellcome Trust Sanger Institute

Genome Center

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the rare assembly was confirmed by restriction digest, except on the occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., MORREP; Information on the MORREP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormrep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl> Rp11-739C15 is from the library RPI1-11.3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

location/Qualifiers

1..110779

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone="Rp11-739C15"

/clone_11b="RPI1-11.3"

/clone_11d="RPI1-11.3"

BASE COUNT 32271 a 25614 c 24906 g 27988 t

ORIGIN

Query Match 19.1%; Score 1189.4; DB 9; Length 110779;

Best Local Similarity 66.9%; Pred. No. 1.6e-276;

Matches 2028; Conservative 0; Mismatches 816; Indels 189; Gaps 16;

QY 1862 TGGCTGAGCTGAGAGTGAACATGCGGCGAGTCTTACAGCCCTCGTGCCTCG 1921

Db 89252 TGGATTTCTTGTGAGAGTGAACAGCTGCTGAGCTCTGAGACCTCGTGCCTCG 89311

QY 1922 GCGCTCTCTGCTGCTGAGCTTCCACTTCTGCTGAGTGAAGAGACCTTACAGCCACCG 1981

Db 89312 GCACCTCTCTGCTGCTGAGCTTCCACTTCTGAGTGAAGAGACCTTACAGCCACCG 89371

QY 1982 TGCACCTGTGGAGACCTTCTGCTGAGTGAAGAGACCTTCTGAGTGAAG 2041

Db 89372 TGCCTGTGGAGACCTTCTGCTGAGTGAAGAGACCTTCTGAGTGAAG 89431

QY 2042 GGGAGTGTGGAGAGAGCTCAAGAGAGACCGGGGCTGCGACGCGCTTGGCGGC 2101

Db 89432 GGGAGTGTGGAGAGAGCTCAAGAGAGACCGGGGCTGCGACGCGCTTGGCGGC 89489

QY 2102 AGCTGAGATTCCGGGTGGGCTTGGCGGGCCCGGCACTGAGAGACGCGGCCAGC 2161

Db 89490 AGCTGAGATTCCAGTGGGCTTGGCGGGCCCGGCACTGAGAGACGCGGCCAGC 89549

QY 2162 CTGTC-----CAGCCCGGGCAATGAGAGCTTACAGCCCGGGCCAGCGCTGCGG 2213

Db 89550 CTGTCAGAGCCCTCGCGCCCGGGCAATGAGAGCTTACAGCCCGGGCCAGCGCTGCGG 89609

QY 2214 AGGCTGACAGTGGGCTGCGCCAGAGTGCACCGCGCGCGCTGCTGCTGCTGCTGCTG 2273

Db 89610 AGGATGACTGAGTGGCTGCGCCAGAGTGCACCGCGCGCTGCTGCTGCTGCTGCTG 89665

QY 2274 ACTGGGCTTACAGAGCTTCCCGCGGGGCGAGGCTGCGGAGCTGAGAGCCCGCATGCTG 2333

Db 89666 ACCGGGCTTACAGTGGCTTCCCGCGGGGCGAGGCTTGGGAGCTGAGAGCCCGCATGCTG 89725

Db 91722 CCAGTGTGGCCCAACAGACACTCTTAATGAGATCTTCAATAACTTAACACT 91781
Qy 4547 TCCGTGTAACCAACCCCTCAAGGCGAGGTGAGTGCATGCTATATATCCACAC 4606
Db 91782 TTCAATCTTAAAGACACCTCTGCGGTGTGTGTGCTCATGCTTATACCTTACAC 91841
Qy 4607 TTTGGAGGCGGAGCGGGGTGATCACTTGAAGTTCAGACCAAGCTTACAC 4666
Db 91842 TTTGGAGGCGGAGCGGGGTGATCACTTGAAGTTCAGACCAAGCTTACAC 91898
Qy 4667 CAACATGTGTAATCCCACTCTTCAATTAAC-AAAATATGATGACATGTGTGCA 4725
Db 91899 CAACATGTGTAATCCCACTCTTCAATTAAC-AAAATATGATGACATGTGTGCA 91958
Qy 4726 TGCTGTATCCCACTCTTGAAGTTCAGACCAAGCTTACACCAAGGAGG 4785
Db 91959 CGCTGTATCCCACTCTTGAAGTTCAGACCAAGCTTACACCAAGGAGG 92018
Qy 4786 GGAAGTTGTAGTGAAGCGAGATGTCGCACTGACCTGAGCAATGAGGAACT 4845
Db 92019 GGAAGTTGTAGTGAAGCGAGATGTCGCACTGACCTGAGCAATGAGGAACT 92078
Qy 4846 CCATCTCAAAAAAACAACCAACCACTC 4878
Db 92079 CCATCTCAAAAAAACAACCAACCACTC 92111

RESULT 14
AL732292 51015 bp DNA linear PRI 19-JUN-2002
LOCUS Human DNA sequence from clone RP4-597N16 on chromosome 1, complete
DEFINITION
AL732292
VERSION AL732292.12 GI:21531517
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 51015)
Lad.H.
Direct Submission
Submitted (19-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail: enquiries:
humquerry@sanger.ac.uk
On Jun 21, 2002 this sequence version replaced gi:21065403.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL, Sw., SwissProt, Tr., TrEMBL, Wp., WormPeP; information on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone comp1 of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP4-597N16 is from the library RP4-4 constructed by the group of
Pleier de Jong, for further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES

Location/Qualifiers
source 1..51015
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP4-597N16"
/clone_11d="RP4-4"
BASE COUNT 14768 a 11872 c 10994 g 13381 t
ORIGIN

Query Match 19.18; Score 1187.8; DB 9; Length 51015;
Best Local Similarity 66.88; Pred. No. 3.3e-276;
Matches 2027; Conservative 0; Mismatches 817; Indels 189; Gaps 16;
Db 1862 TGCGTTGAGCGTGAAGTGAAGCATGCGGAGTCTTCAAGACCCCTGCTGCTCG 1921
Qy 46767 TGGATTCTTGTGAGAGTGAAGCATGCTGCTGCTTCAAGACCCCTGCTGCTCG 46826
Db 1922 GCGCTTCTTGTGCTGCGCTTCAAGTGTGCTGCTTGAAGACCCCTTGAAGCCGCG 1981
Qy 46827 GCACCTCTTGTGCTGCGCTTCAAGTGTGCTGCTTGAAGACCCCTTGAAGCCGCG 46886
Qy 1982 TGCACTGTGAGAGCCCTTCTGCGCTGCGCAAGCCGAGCCGCTGCTGCTGCA 2041
Db 46887 TGCGCTGTGAGAGCCCTTCTGCGCTGCGCAAGCCGAGCCGCTGCTGCTGCA 46946
Qy 2042 GGAAGTGTGAGAGAGAGCTCAAGCAGAACCGGGGCTGCGCAGCGGCTTGCAGGCG 2101
Db 46947 GGAAGTGTGAGAGAGAGAGCTCAAGCAGAACCGGGGCTGCGCAGCGGCTTGCAGGCG 47004
Qy 2102 AGCTGAGTTCCGGGTGGGCTGCGCTTGGCGGCGCCGCACTGAGACAGCGGCGCAGC 2161
Db 47005 AGCTGAGTTCCAGAGTGGGCTGCGCTTGGCGGCGCCGCACTGAGACAGCGGCGCAGC 47064
Qy 2162 CCGTGC-----CAGGCGCGGCGCAATGAGAGCTTGAAGACCGGCGCGGCTGCGG 2213
Db 47065 CCGTGCAGCCCTGCGGCGCGGCGCAATGAGAGCTTGAAGACCGGCGCGGCTGCGG 47124
Qy 2214 AGGCTGTACTGGGTGCGGCGCAAGTGGCAGCCCGCGGCTGCTGCTGCTGCTGCTGCTG 2273
Db 47125 AGGATGTACTGGGTGCGGCGCAAGTGGCAGCCCGCGGCTGCTGCTGCTGCTGCTGCTG 47180
Qy 2274 ACTGGGCTTGAAGAGCTTCCCGGCGGCGAGGCTGCGGACCTGACAGCCCGCATGCGCTG 2333
Db 47181 ACCGGGCTTGAAGAGCTTCCCGGCGGCGAGGCTGCGGACCTGACAGCCCGCATGCGCTG 47240
Qy 2334 AGCT-----CCCTTCATGGGCTCTGTGGGCGGCGAGGCTCCCGAGCAGCAGCC 2388
Db 47241 AGCTCCAGCCCTTCCATGGGCTCTGTGGGCGGCGAGGCTCCCGAGCAGCAGCC 47300
Qy 2389 CCGTCTCAGAGCGCGCCAGTCCATGCAACGAGGAGGCTGAGAGTGGGCGCAGG 2448
Db 47301 CCGTCTCAGAGCGCGCCAGTCCATGCAACGAGGAGGCTGAGAGTGGGCGCAGG 47360
Qy 2449 CACCGGAGTGGCAGGAGTACCCCTGCAAGCCCTGCTGGGAAATTCATGCTGGGGAAGCC 2508
Db 47361 CA-TGGGAAATGGCAGGAGTCCACCTGCAAGCCCGGCTGAGGATTCATGCTGGGGAAGCC 47419
Qy 2509 AGCTGGCTCTGAGTGTGGGAGTGGAGAACTTATGCTGCTGAGGAGTGT 2568
Db 47420 AGCTGGCTCTGAGTGTGGGAGTGGAGAACTTATGCTGCTGAGGAGTGT 47479
Qy 2569 AAATACACATACAGACCCCTGCTGAGTCAAGGCTGCTGATGACCAATTCACACT 2628
Db 47480 AAATACACATACAGACCCCTGCTGAGTCAAGGCTGCTGATGACCAATTCACACT 47539

Db	Accession	Source	Organism	Reference Authors	Title	Journal	Reference Authors
Db	49534	GGAGTTGGCAATGAGCGGGATCGACACACTGCACCTCGACCGCTGTGACAGAGAGAGT	4953				
Qy	4846	CCATCTCAAAAAACACACAAAAACCACCTC	4878				
Db	49594	CCATCTCAAAAAACAAAAAAGACAC	49626				
RESULT 15							
AC104212							
LOCUS	AC104212	166806 bp	DNA	linear	PRI 27-APR-2002		
DEFINITION	Homo sapiens chromosome 8, clone RP11-941H19, complete sequence.						
ACCESSION	AC104212						
VERSION	AC104212.3	GI:20149422					
KEYWORDS	HNC						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 166806)						
AUTHORS	Barren, B., Linton, L., Nusbaum, C. and Lander, E.						
TITLE	Homo sapiens chromosome 8, clone RP11-941H19						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 166806)						
AUTHORS	Barren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, J., Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamaczars, R., Landers, T., Lehoczy, J., Levine, N., Matthews, C., Maclean, C., Macdonald, P., Major, J., Margus, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhhang, P., Plerre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Sclaus, N., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tophan, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						
TITLE	Direct Submission						
JOURNAL	Submitted (30-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA						
REFERENCE	3 (bases 1 to 166806)						
AUTHORS	Barren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, J., Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamaczars, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margus, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoli, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhhang, P., Plerre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Sclaus, N., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tophan, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						
TITLE	Direct Submission						
JOURNAL	Submitted (15-APR-2002) Whitehead Institute/MIT Center for Genome						

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Research, 320 Charles Street, Cambridge MA 02141, USA 4 (bases 1 to 196606)	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, K., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karates, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Leticosky, J., Levine, R., Linblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKean, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, O., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K. C., Phunkhang, P., Pierrel, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rossetti, M., Roy, A., Santos, R., Schauer, S., Schnapack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	Submitted (27-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Apr 15, 2002 this sequence version replaced g1.2018254. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RN/RepeatMasker.html		
Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence.submissions@genome.wi.mit.edu	Project Information Center project name: L19698 Center clone name: 94L_H19	Location/Qualifiers 1. 196806 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="8" /map="8" /clone="RP11-941H19" /clone.lib="RPC1-11 Human Male BAC" 641. 823 /rpt_family="L1ME4A" 896. 1076 /rpt_family="MER5A" 1804. 1848 /rpt_family="(CA)n" complement(2152..2310) /rpt_family="AluUo" 2311. 2338 /rpt_family="(TG)n" complement(2339..2455) /rpt_family="AluUo" 3098. 3469 /rpt_family="THE1B" 3470. 4964 /rpt_family="THE1B-Int" 4966. 5330 /rpt_family="THE1B" 5420. 5858 /rpt_family="L1R54B" 6492. 6791 /rpt_family="AluX" 8696. 8746 /rpt_family="(CA)n"		

repeat_region 8887..8943
/rpl_family="Tigger1"
complement(8981..9135)
repeat_region /rpl_family="Alus/x"
9136..9303
repeat_region /rpl_family="Tigger1"
9595..9906
repeat_region /rpl_family="Alusp"
complement(9979..10503)
/rpl_family="MIR1G"
repeat_region 10839..10909
/rpl_family="(GAGTG)n"
12234..12305
repeat_region /rpl_family="(CACCC)n"
12402..12480
repeat_region /rpl_family="L2"
12560..12649
repeat_region /rpl_family="AluDb"
12653..12964
repeat_region /rpl_family="AluYa5"
complement(13206..13302)
/rpl_family="MIR"
repeat_region 13462..13709
/rpl_family="MIR"
repeat_region 13904..14194
/rpl_family="AluDb"
14763..15084
repeat_region /rpl_family="Alusx"
15706..15726
repeat_region /rpl_family="(TA)n"
complement(15727..16021)
/rpl_family="Alusx"
16505..16647
repeat_region /rpl_family="FRAM"
16649..16771
repeat_region /rpl_family="L1MB2"
16774..16959
repeat_region /rpl_family="FRAM"
16961..17635
repeat_region /rpl_family="L1MB2"
18007..18097
/rpl_family="Ga-rich"
18360..18517
repeat_region /rpl_family="L1P"
18595..18920
repeat_region /rpl_family="Alusx"
19031..19319
repeat_region /rpl_family="L1MC/D"
19421..19909
repeat_region /rpl_family="MER34"
20029..20299
repeat_region /rpl_family="Alusx"
complement(20300..20524)
/rpl_family="Alusp"
20371
unsure /note="probably A; possibly G"
20615..20915
repeat_region /rpl_family="Alusx"
complement(22970..23273)
repeat_region /rpl_family="Alusx"
23933..24001
repeat_region /rpl_family="(TA)n"
24126..24422

Query Match 18.1%: Score 1129.8; DB 9; Length 196806;
Best Local Similarity 69.6%: Pred.No.4.9e-262;
Matches 1825; Conservative 0; Mismatches 652; Indels 144; Gaps 16;

QY 1807 GACCAAGCCTTACACTGTAAATCCAGCACTTTGGAGGCAAGCGGGTGAGTGCT 1866
DB 85566 GGCACAGGCTCATGTCATGATCCAGCACTTTGGAGGCGGAAATGGGAG----- 85618
QY 1867 TGACCTTGAGAGTGACAGCATGGCGGAGTCTGCTGCTGCTGCGGCC 1926

DB 85619 -----TGAGAGGTGACAGTGTGTCGAGAGTCTCAGAGCCCTCTGCTGCGGCT 85672
QY 1927 TCCTTGCCCTGGGCTCCACATTCGGTGGCACTTTGAGAGACCTTACGCCACCGCTGCAC 1986
DB 85673 TCCTTGCCCTGGGCTCCACATTCGGTGGCACTTTGAGAGACCTTACGCCACCGCTGCAC 85732
QY 1987 TGTGGAGCCCTTCTGCGGCTGGGCAAGGCCAGAGCCGCGCTCCCTGACCTTGACGGGAG 2046
DB 85733 TGTGGAGCCCTTCTGCGGCTGGGCAAGGCCAGAGCCGCGCTCCCTGACCTTGACGGGAG 85792
QY 2047 GTGTGAGAGGAGAGGCTCAAGCAGAGAACCGGGCTGCGACCGGCTTGCGGCCAGCTG 2106
DB 85793 GTGTGAGAGGAGAGGCGGAGAGAGAACCGGGCTGCGACCGGCTTGCGGCCAGCTG 85852
QY 2107 GAGTTCGGGCTGGGCTGGGCTTTGGCGGGCCCGCACTGAGACAGCGGCCACCGCTGCG 2166
DB 85853 GAGTTCGGGCTGGGCTGGGCTTTGGCGGGCCCGCACTGAGAGCGGCCCGCTGCG 85912
QY 2167 CAGGCCCGGCAATGAGAGGCTTAGCACCGGGGCGAGCGGCTGGGAGAGGTGTACTGG 2226
DB 85913 C-GGCCCGGCAATGAGAGGCTTAGCACCGGGGCGAGCGGCTGGGAGAGGTGTACTGG 85971
QY 2227 TGCCCCAGCAGTGCAGCGCCCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2286
DB 85972 TGCCCCAGCAGTGCAGCGCCCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 86027
QY 2287 ACCCTTCCCGGGGCGAGGCTGCGGAGCTGCAAGCCCGCATGCTGAGCCT-CCC 2341
DB 86028 TCCCTTCCCGGGGCGAGGCTGCGGAGCTGCAAGCCCGCATGCTGAGCCTTCCCGG 86086
QY 2342 CTCGATGGGCTGCTGCGGCGGCGGAGCCCTCCCGAGAGGACAC-ACCCCTGCTCAGAG 2400
DB 86087 CTCGATGGGCTGCTGCGGCGGCGGAGCCCTCCCGAGAGGACAC-ACCCCTGCTCAGAG 86146
QY 2401 CCCCCAGTCCATTCGACGACGCAAGGGCTGAGAACTGGGGGCGCAAGCGGACGCACTGG 2460
DB 86147 CCCCCAGTCCATTCGACGACGCAAGGGCTGAGAACTGGGGGCGCAAGCGGACGCACTGG 86205
QY 2461 CAGGCAAGTACCCCTGCAAGCCCTGCTGGGAAATCCACTGGGTGAAGCCAGCTGGGCTCT 2520
DB 86206 CAGGCAAGTACCCCTGCAAGCCCTGCTGGGAAATCCACTGGGTGAAGCCAGCTGGGCTCT 86265
QY 2521 GAGTCTGGTGGAGACTTGGAGAACTTATGCTAGCTCAGGATGCTGAATACCAAT 2580
DB 86266 GAGTCTGGTGGAGACTTGGAGAACTTATGCTAGCTCAGGATGCTGAATACCAAT 86325
QY 2581 CAGCAAGCTGTGTAGCTCAGGAGTCTGTGAATGACCAATCAGACTGTATAGCT 2640
DB 86326 CAGCAAGCTGTGTAGCTCAGGAGTCTGTGAATGACCAATCAGACTGTATAGCT 86385
QY 2641 ACTGTGAGGGGCTTGGAGAACTTATGCTAGCTCAGGATGCTGAATACCAATC 2700
DB 86386 GCTGTGAGGGGCTTGGAGAACTTATGCTAGCTCAGGATGCTGAATACCAATC 86445
QY 2701 GGCACCTGTATCTAGCTCAGGAGTCTGTGAATACCAATCAGCCCTGTATAGCT 2760
DB 86446 AGCACCCTGTGTAGCTCAGGAGTCTGTGAATACCAATCAGCCCTGTATAGCT 86505
QY 2761 AGGATGTGAGTGAATGACCAATGACAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTG 2820
DB 86506 CTCGTGTGGGGCTTGGAGAACTTATGCTAGCTCAGGAGTGTGAATACCAATCAGCTA 86565
QY 2821 GAAGAGACCAACAGGCTTGTGTGAGCAATTAAGCTTCTACCTGCTGCTGCTGCTGCTG 2880
DB 86566 GCACCCCTGTGTAGCTCAGGAGTGTGTAG-----TGACCAATCAACACTGTATAGCT 86619
QY 2881 GGGCTGAGTCCGAAGAAGAGAGTCAAGGAGAGTAAGGCTGGGCGCTTGTATAGAT 2940
DB 86620 -----AGTGTCTGTGGGGCTTGTGAGAACTT 86650
QY 2941 TTGGGTAGTAAGGAATTTACAGTCAAGGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3000

Db 86651 ATGTCAGCTCAGGGATTGTAAACACACTAATCAGACCCCTGTGTTAGCTCAGGGTTTG 86710
QY 3001 GGGGTCCGAAGTCTCTAGTGGGGTGTCTTTTATGACCGAGATATAGCCAGGAAAAAGCACT 3060
Db 86711 TGAATGACCAANTGACACTCTGTAT-CTAGCTCTCTGTGGGGCTTGGAGAACCTTT 86769
QY 3061 TTCACAGTAATGTCAATCAATTAAGCAAGGACCCGCATTTACACTCTTTTGTGGTG 3120
Db 86770 ATGTCATCTCAGGAGATTGTAAATACACAAATCAGCACTCTGTATCTAGCCCAAGTTTG 86829
QY 3121 GAATGTATCAGTTAAGTTGGGGCAGGGCATATTCACCTCTTTTGTGATTCCTTCAGTTAC 3180
Db 86830 TAAACACCAATCA-----GCAACCTGTGTTTAA 86859
QY 3181 TTCAGGCCATCTGGGCGCTATATGTGCAAGTTACAGGGAGTCCATGGCTTGGGCT 3240
Db 86860 CTCAGGGTTGTGTGTCACCAATTA----- 86886
QY 3241 CAGAGCGTTGACACTACTCTGTGGGGGCTTGGAGAAATGTTGTGCACTCTGTAT 3300
Db 86887 CACTGTATCTAGCTCTCTGTGGGGCTTGGAGAACTTTGTGTGATCTCTGTAT 86946
QY 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTTGTCTAGCTCAGGAGATTGTAAACGCA 3360
Db 86947 CTAACTAATCTGATGGGAGCTGGAGAACAGTTATATCTAGCTCAGGAGATTGTAAAGCA 87006
QY 3361 CCATACGCGCCCTGTGTAACAAACAGACACCTGGCTCTACCAATAGACAGATGTGGGTGG 3420
Db 87007 CCATACGACCCCTGTGTAACAAATAGCACTGGCTCTACCAATAGTGGGTGG 87066
QY 3421 GGCCAGATTAAGAGAAATAAAGCAGGCTGCCGAGCCAGAGTGGCAACGGCAGACAGTCC 3480
Db 87067 GGCCAGATTAAGAGAAATAAAGCAGGCTGCCGAGCCAGAGTGGCAACCCGCTGGGTCC 87126
QY 3481 CTATCCAAATATGGCAGCTTTGTTCTTTTGTGTTGCGATTAATCTTGTACTGCTCG 3540
Db 87127 CTTTCCAGCTGTGGAGCTGTGTTCTTTGTTCTTTCATTAATCTTGTGCTGCTCA 87186
QY 3541 CTTTGTGGTCCACAGCTCTTTATGAGC-----TGTAACTCACCAGAACTCTGAC 3596
Db 87187 GTCTTGTGGTCCACAGCTCTTTATGAGCTGTATGATTAACATCTCACCGGAAGATCTCAG 87246
QY 3597 CTTTACTCTCTGAAGCCACTAAGACACAGACCCACCGGAGGAATGAACAACCTCCGGCG 3656
Db 87247 CTTTACTCTCTGAAGCCACGAGACACAGACCCACCGGAGGAATGAACAACCTCCAGACA 87306
QY 3657 CGCTGCTTAAGAGCTTAACACTCACGCGGAAGTCTGCACTTCACTCTCCAGCCAGC 3716
Db 87307 CGCTGCTTAAGAGCTTAACACTCTCACGCGGAAGTCTGCACTTCACTCTCCAGCCAGT 87366
QY 3717 GAGACCAAGAACCCACACAGAGAAAGAACTGGGAACACATCTGACATCAGAAAGACA 3776
Db 87367 GAGACTAGCAACCCACAGAGAAAGAACTCCGAAACATCTGAAACAGCAGAAAGGAAA 87426
QY 3777 AACTCAGATGACACACTTAAGAGCTTAACACTCAGAGGGTCCGGGCTTCTT 3836
Db 87427 GACTCCAGATGACACCGCTTAAGAGCTTAACAGGCGCGCGCGT---GGCTCACGCC 87483
QY 3837 CTTGAAGTCAGTAGAACCAAGCACTCACAGTTTCGACACAAAGCCAGAGATTGAGAT 3896
Db 87484 TGTAAACCCGCACTTTGGAGGCCGAGGCGGCGATCACAGAGTCAGAGATTCAGAC 87543
QY 3897 CAGCTGGGGCAATATGATAATGCGCTCTCTGCAAAAAAATAAATAATTAACAATAATG 3956
Db 87544 CATCCCGGCTAAAGCGGTAAACCCGCTCTCTACTAAAAATTA-----AAAAATTA 87595
QY 3957 GCGAGCATGTGTGCTGCTGTGTGCCACACTACGCGGAGGCTAAAGTGGAGAT 4016
Db 87596 GCGGGGCTAGTGGGCGGCGCTGTAGTCCACTACTTGGAGGCTGAGGCGAGAGAAAT 87655
QY 4017 CGCTTAAGCTGGGAGGTGAAGACTGAGCTGTGATTTGACCAACAGCCCTTAAGGC 4076
Db 87656 GGGGTGAACCGGAGGCGGAGCTTGACGTAGACCGAGATCCGCACTGACTCCAGCC 87715

QY 4077 TGGGGGACAGACTAGACCCCTGTTCCCTCCGCAAAAAATTGACAAA---GTGTAA 4132
Db 87716 TGGAGACAAAGCCAGACTCCGCTCAAAAAAAAAAAAAAAAAAAGCTGTAA 87775
QY 4133 TAAAGGGGCTGATATGAGGCTAAGGCGAGTGGCTCATGGCTTAATCCAGCACTTTGGG 4192
Db 87776 CACTCATGTCCAGGCTCGCGGCTTTCTTTTGAAGTACGTAGAACCAAGAACCCAGAA 87835
QY 4193 AAGCCAGGCGGCGGCTCACTAAGTCAAGAGTGTAGAACCAAGCTTGGCCAAATGGA 4252
Db 87836 TTCGGAACCCAGAGGATGTGTGAACGTAGAGATTGAGACCAAGCTTGGGCAACAAAGT 87895
QY 4253 GAAAGCCATCTCTTAAATAATACAAATTAAGCCGCTGTGGGGCAGTGTGAGCAT 4312
Db 87896 ATAGTAAGACCTCTTCTGCAAAAAATGAATTAAGTGGGTGTGTGTGTC 87955
QY 4313 GCTGTATATCCACTACTCTCAGAGGCTGAGGCGAGAGAAATCACTTAACCCAGAGCGC 4372
Db 87956 ACATGTATGCCACACTACTGTGGGAAAGCTGAGGTGGCAGAAATTACTTAACCCAGAGGTC 88015
QY 4373 GCGGTTCAGTAGACCGAGATCTGCAATTGCACTCCACC 4413
Db 88016 AAGGCTCAGTAGACCCCTGGCAGCACCGCTGTAGTCCAGCC 88056

Search completed: August 16, 2003, 20:10:10
Job time : 14696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 22:51:15 ; Search time 998 Seconds
(without alignments)
16864.726 Million cell updates/sec

Title: US-09-845-020A-5
Perfect score: 6235
Sequence: 1 gatcacttgaggacagtagt.....tgcacttccccacagagcct 6235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6235	100.0	6235	21	AAZ29169 Human G-CSF genomic
2	6235	100.0	6679	21	AAZ29170 Partial sequence o
3	6235	100.0	6679	22	AAZ29170 Human granulocyte
4	3725	59.7	6753	22	AAZ29173 Human interferon a
5	2834	45.5	2834	21	AAZ29173 Targeting sequence
6	2220.4	35.6	2515	22	AAZ29174 Human granulocyte
7	1252	20.1	1252	21	AAZ29174 Targeting sequence
8	773.4	12.4	50000	21	AAA96363 Polymorphic repeat

C	9	757.4	12.1	163350	24	AAZ46127	Human tumour suppr
C	10	744.8	11.9	260209	24	AB555564	Human SUF2 genomic
	11	733.2	11.8	121724	24	ABQ88143	Human osteoblast d
	12	710.8	11.4	49959	20	AAZ23901	Human LOBO homolog
	13	708.4	11.4	55235	22	AAZ67426	Human immune/haema
	14	693.8	11.1	22081	22	AAZ97862	Human neuroblastom
	15	669.2	10.7	1831	21	AAA27481	5'HSS endogenous r
	16	668	10.7	185371	24	ABT10718	Human breast cance
C	17	652.8	10.5	10351	24	ABK86932	Human CDLE cDNA.
C	18	614.2	9.9	5236	22	AAK84744	Human immune/haema
C	19	608.6	9.8	2610	22	AAK94741	Human full-length
C	20	606.2	9.7	19974	24	AB555025	Invertebrate forag
C	21	593.6	9.5	1914	24	ABX17314	Human cancer promo
C	22	593	9.5	249899	25	ABZ808229	Human transdrin ge
	23	592.6	9.5	122888	24	ABK83569	Human cDNA differe
C	24	584.8	9.4	1872	23	AA583445	DNA encoding novel
C	25	583.8	9.4	6766	22	AAK78681	Human immune/haema
C	26	583.8	9.4	6766	23	ABK42991	Genomic sequence #
C	27	583.2	9.4	2597	22	AAH15876	Human cDNA sequenc
C	28	576.6	9.2	1405	23	AA590241	DNA encoding novel
C	29	573	9.2	1474	23	AA578897	DNA encoding novel
C	30	564.4	9.1	28818	22	AAJ35901	Human musculoskele
C	31	564.4	9.1	28818	25	ABX58889	CDNA encoding nove
C	32	563.8	9.0	9704	24	ABL60777	Genomic DNA of an
C	33	563.8	9.0	9704	25	ABX93511	Human aminocyclase
C	34	560.4	9.0	32127	22	AAI99255	Human excretory re
C	35	560.4	9.0	32127	22	AAI63605	Human kidney relat
C	36	560.4	9.0	72215	22	AAK86832	Human immune/haema
	37	560	9.0	1503841	24	ABT00010	Human neuregulin 1
	38	560	9.0	1503841	24	ABT01503	Human neuregulin 1
	39	560	9.0	1503900	22	AAK95240	Human neuregulin-1
	40	560	9.0	1503900	22	AAK96733	Human neuregulin-1
	41	551.8	8.9	3943	23	AA580270	DNA encoding novel
	42	549.2	8.8	3167	22	AAH95222	Human protein enco
C	43	546	8.8	1349	23	AA589767	DNA encoding novel
C	44	542.8	8.7	14273	22	ABA07392	Human pancreatic c
	45	542.8	8.7	14273	22	AAK91123	Human digestive sy

ALIGNMENTS

RESULT 1
AAZ29169 standard: DNA: 6235 BP.
AC AAZ29169;
DT 21-FEB-2000 (first entry)
DE Human G-CSF genomic sequence upstream of transcription start site.
XX Granulocyte colony stimulating factor; G-CSF; genomic sequence;
XX transcription start site; 5' non-coding sequence; DNA construct;
XX targeting sequence; regulatory region; marker gene; selection;
XX homologous recombination; gene therapy; delivery system; upstream;
XX hematopoietic progenitor cell; chemotherapy-induced neutropenia;
XX bone marrow transplantation; congenital neutropenic disorder;
XX chronic idiopathic disorder; ds.
OS Homo sapiens.
XX
XX W09957291-A1.
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-US09924.
XX
XX 07-MAY-1998; 98US-0084649.
XX
XX (TRAN-) TRANSKARYOTIC THERAPIES INC.
XX
XX Treco DA, Heartlein MW, Selden RF;

XX WPI: 2000-072235/06.

XX Novel genomic sequences used for treating human diseases and disorders

PS Claim 1; Fig 5; 58pp; English.

CC The present DNA sequence is the human genomic sequence from a region
CC upstream of the transcription start site of granulocyte colony-
CC stimulating factor (G-CSF). This sequence represents nucleotides -6578
CC to -364 relative to the translation start site. G-CSF gene contains five
CC exons and four introns. A DNA construct comprising a targeting sequence
CC homologous to the 5' non-coding sequence of G-CSF, a transcriptional
CC regulatory sequence that differs from the endogenous G-CSF gene, and a
CC selectable marker gene for selection of recombinant cells is generated.
CC It can be used by homologous recombination to modify the expression of
CC endogenous G-CSF. These recombinant cells which express G-CSF are useful
CC for in vitro production of the protein and gene therapy. Such cells may
CC also be used in a delivery system for stimulating the proliferation and
CC differentiation of hematopoietic progenitor cells, or for other
CC conditions that can be treated with G-CSF, like chemotherapy-induced
CC neutropenia, to treat patients undergoing bone marrow transplantation,
CC chronic idiopathic and congenital neutropenic disorders.

XX Sequence 6235 BP; 1586 A; 1610 C; 1708 G; 1331 T; 0 other;

Query Match 100.0%; Score 6235; DB 21; Length 6235;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGCTGAGGACAGTTCAGTTCAGACACCTGGGAGCATGAGGAGACTGTCTACG 60
DB 1 GATCAGCTGAGGACAGTTCAGTTCAGACACCTGGGAGCATGAGGAGACTGTCTACG 60
QY 61 AAAATCAAAAAATTTATGGCGGCGATGTCGTCAGTTCATCCGAACTTGGG 120
DB 61 AAAATCAAAAAATTTATGGCGGCGATGTCGTCAGTTCATCCGAACTTGGG 120
QY 121 ACATCAAGGCAAGTGCATCTGAGGTCAGAGTTCAGACTAGCTGGCCAACTGTT 180
DB 121 ACATCAAGGCAAGTGCATCTGAGGTCAGAGTTCAGACTAGCTGGCCAACTGTT 180
QY 181 GAAACCTTATCTCCACTTAAAAATTAAGCAGCATGTCGTCGAGCAGCTGTA 240
DB 181 GAAACCTTATCTCCACTTAAAAATTAAGCAGCATGTCGTCGAGCAGCTGTA 240
QY 241 ATCCGCGCTACTCAGAGGCTGAGGAGAGATCACTTGAACCCAGAGGCGAGGTTG 300
DB 241 ATCCGCGCTACTCAGAGGCTGAGGAGAGATCACTTGAACCCAGAGGCGAGGTTG 300
QY 301 CAGTGAAGCTGAGATCAACACCACTGCAGCTGGGTGACAGAGCAAGACTCTATCTC 360
DB 301 CAGTGAAGCTGAGATCAACACCACTGCAGCTGGGTGACAGAGCAAGACTCTATCTC 360
QY 361 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
DB 361 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
QY 421 CTACTCAGAGAGGCTGAGGAGATCACTTGAACCTGGGAGAGTCAAGGCTTAAAG 480
DB 421 CTACTCAGAGAGGCTGAGGAGATCACTTGAACCTGGGAGAGTCAAGGCTTAAAG 480
QY 481 GCCAAGATCTGCACCTACCTCAGCTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 GCCAAGATCTGCACCTACCTCAGCTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
DB 541 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 ATGTATATAGTTTGAACCTCAAGATCCAGATAGTCAATTTTATGGGCTTGGGCGCT 660
DB 601 ATGTATATAGTTTGAACCTCAAGATCCAGATAGTCAATTTTATGGGCTTGGGCGCT 660

DB 601 ATGTATATAGTTTGAACCTCAAGATCCAGATAGTCAATTTTATGGGCTTGGGCGCT 660
QY 661 ATGTCTCTGTCAACATCACTGCTGCTCTTCTAGACAAAGCAGCTATTAACAT 720
DB 661 ATGTCTCTGTCAACATCACTGCTGCTCTTCTAGACAAAGCAGCTATTAACAT 720
QY 721 ACATACATGAATTTTATAGACATGAGATTTGAATTTATGATTTTATACATTTTAT 780
DB 721 ACATACATGAATTTTATAGACATGAGATTTGAATTTATGATTTTATACATTTTAT 780
QY 781 AAAATATCTTTTAAAAATTTTCCCTAACCATTTTAAAAAGTAAAAAGCGGCACGCG 840
DB 781 AAAATATCTTTTAAAAATTTTCCCTAACCATTTTAAAAAGTAAAAAGCGGCACGCG 840
QY 841 GCCATCTGACGCTGTAAATTCAGACCTTTGGAGGCTGAGGAGATCATCTTGAAG 900
DB 841 GCCATCTGACGCTGTAAATTCAGACCTTTGGAGGCTGAGGAGATCATCTTGAAG 900
QY 901 ATCAACGTTTCGAGACAGGCTGGCCAACTAGCAAAACCCCATTTCTACTAAAAATTA 960
DB 901 ATCAACGTTTCGAGACAGGCTGGCCAACTAGCAAAACCCCATTTCTACTAAAAATTA 960
QY 961 AAAATTAAGTGGGATAGTGTGTCACACCTGTATCCACCTTCTGGAGGCTGAGGCA 1020
DB 961 AAAATTAAGTGGGATAGTGTGTCACACCTGTATCCACCTTCTGGAGGCTGAGGCA 1020
QY 1021 GGAGAAATCGCTTGAACCTGGGAAAGGAGGTTGAGTGCAGTGCACCAATCATGTCAC 1080
DB 1021 GGAGAAATCGCTTGAACCTGGGAAAGGAGGTTGAGTGCAGTGCACCAATCATGTCAC 1080
QY 1081 TCCAGCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTAAAGCCAT 1140
DB 1081 TCCAGCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTAAAGCCAT 1140
QY 1141 TCCATTAATTCAGTACATCACTGATACATCAAGTGTGGTACTGCTGTCGAGGCT 1200
DB 1141 TCCATTAATTCAGTACATCACTGATACATCAAGTGTGGTACTGCTGTCGAGGCT 1200
QY 1201 ACCGTGAGAGTAGAGTCTTGTGTCACAGAGATACATTTCCATTAATTAAGACT 1260
DB 1201 ACCGTGAGAGTAGAGTCTTGTGTCACAGAGATACATTTCCATTAATTAAGACT 1260
QY 1261 ACCAAGTTCATTCAGAGAGGTTTTTTTTTACATCTACCTCCCCAGCAACAAT 1320
DB 1261 ACCAAGTTCATTCAGAGAGGTTTTTTTTTACATCTACCTCCCCAGCAACAAT 1320
QY 1321 GAGAGTACTCCAGATCTTTTACAAAGATGCTTAAGCCAGTACAGATGAAGAAACGGA 1380
DB 1321 GAGAGTACTCCAGATCTTTTACAAAGATGCTTAAGCCAGTACAGATGAAGAAACGGA 1380
QY 1381 AGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 AGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 GATAGCTGGAAGATGAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 GATAGCTGGAAGATGAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 ACTTCTGAGGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1560
DB 1501 ACTTCTGAGGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1560
QY 1561 CACTCTCTGTTTGTCTTATTTCTCCCATGTTGGGCTGAAGTCTGATTTGAGCCGTTAT 1620
DB 1561 CACTCTCTGTTTGTCTTATTTCTCCCATGTTGGGCTGAAGTCTGATTTGAGCCGTTAT 1620
QY 1621 TCAAGATGTCACGCTTTCTTGACAGAACTAGTTCACAGAAACAGAGGAGGCTTGGA 1680
DB 1621 TCAAGATGTCACGCTTTCTTGACAGAACTAGTTCACAGAAACAGAGGAGGCTTGGA 1680
QY 1681 AGATGATCTTAACCTCAAGATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1740
DB 1681 AGATGATCTTAACCTCAAGATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1740

QY	1741	AACTTTTTTCACTTCTCTGAGGCCATCCCTTGGCTACAAACACCACTGTGTTGATACAGGA	1800
Db	1741	AAATTTTTTCACTTCTCTGAGGCCATCCCTTGGCTACAAACACCACTGTGTTGATACAGGA	1800
QY	1801	TGAATATGACGAAGTCCCTTACCACTGTATATCCAGCACTTTGGAGAGGCGAGGGGGTGG	1860
Db	1801	TGAATATGACGAAGTCCCTTACCACTGTATATCCAGCACTTTGGAGAGGCGAGGGGGTGG	1860
QY	1861	ATGGCTTGAGCCTGAGAGGTACAGCATGCGGAGAGTCTCACAGCCCTCTGTGCCTTC	1920
Db	1861	ATGGCTTGAGCCTGAGAGGTACAGCATGCGGAGAGTCTCACAGCCCTCTGTGCCTTC	1920
QY	1921	GGCGGCTCTCTGGCTGGGGCTCCACATTCCGTTGGTCACTTGGAGAGCCCTTACGCCACCG	1980
Db	1921	GGCGGCTCTCTGGCTGGGGCTCCACATTCCGTTGGTCACTTGGAGAGCCCTTACGCCACCG	1980
QY	1981	CTGCACGTGTGGAGCCCTCTTCTGGGCTGGCCAGAGCCAGACCGGCTCCCTTCACACTTGC	2040
Db	1981	CTGCACGTGTGGAGCCCTCTTCTGGGCTGGCCAGAGCCAGACCGGCTCCCTTCACACTTGC	2040
QY	2041	AGGAGAGGTGTGGAGGAGAGGCTCAACAGGAACCCGGGCTGCGCAGCGGCGCTTGGGGGC	2100
Db	2041	AGGAGAGGTGTGGAGGAGAGGCTCAACAGGAACCCGGGCTGCGCAGCGGCGCTTGGGGGC	2100
QY	2101	CAGCTGGAGTTCCGGGTGGGGGTGTGGCGGGCCCGCAGCTCTGAGCAGCGGGCGAG	2160
Db	2101	CAGCTGGAGTTCCGGGTGGGGGTGTGGCGGGCCCGCAGCTCTGAGCAGCGGGCGAG	2160
QY	2161	CCCTGGCCAGGCCCCGGGGCAATGABAAGCTTATAGCACCGGGCCAGCGGCTGCGGAGGTGT	2220
Db	2161	CCCTGGCCAGGCCCCGGGGCAATGABAAGCTTATAGCACCGGGCCAGCGGCTGCGGAGGTGT	2220
QY	2221	ACTGGGTGCCCCAGCAGAGTGCAGGCCCGCGGCGGTGTGCTCTGCATTTTCTCACTGGGC	2280
Db	2221	ACTGGGTGCCCCAGCAGAGTGCAGGCCCGCGGCGGTGTGCTCTGCATTTTCTCACTGGGC	2280
QY	2281	CTTATGACAGCCTTCCCGGGGGGCGAGGGCTCGGGACCTGACGCGGCCATGCTGAGCCTTC	2340
Db	2281	CTTATGACAGCCTTCCCGGGGGGCGAGGGCTCGGGACCTGACGCGGCCATGCTGAGCCTTC	2340
QY	2341	CCTCCATGGGGTCCGTTGTCGGGCGGAGCGTCCCGAGAGAGCAACCCCTGTGCTCCACAG	2400
Db	2341	CCTCCATGGGGTCCGTTGTCGGGCGGAGCGTCCCGAGAGAGCAACCCCTGTGCTCCACAG	2400
QY	2401	CGCCCACTCCCATCGACACGACGCAAGGGCTGAGAAATGCGGGCGACGACCGGAGACTGG	2460
Db	2401	CGCCCACTCCCATCGACACGACGCAAGGGCTGAGAAATGCGGGCGACGACCGGAGACTGG	2460
QY	2461	CAGGACAGTACCCCTGCAAGCCCTGTGTGGGAAATCCACTGGGTGAAGCACTGTGGGCTCT	2520
Db	2461	CAGGACAGTACCCCTGCAAGCCCTGTGTGGGAAATCCACTGGGTGAAGCACTGTGGGCTCT	2520
QY	2521	GAGCTGTGGTGAAGACTTGGAGAACCTTATATCTAGCTCAGGAGATCGTAATTAACCAAT	2580
Db	2521	GAGCTGTGGTGAAGACTTGGAGAACCTTATATCTAGCTCAGGAGATCGTAATTAACCAAT	2580
QY	2581	CAGCACCCCTGTCTAGCTCAGGGTCTGTGTGANTCCACCAATCCACACTCTGTATCTAGCT	2640
Db	2581	CAGCACCCCTGTCTAGCTCAGGGTCTGTGTGANTCCACCAATCCACACTCTGTATCTAGCT	2640
QY	2641	ACTCTGATAGGGGCGCTTGGAGAACCTTATATGTCTAGCTCAGGGATTTGTAATTAACCAATC	2700
Db	2641	ACTCTGATAGGGGCGCTTGGAGAACCTTATATGTCTAGCTCAGGGATTTGTAATTAACCAATC	2700
QY	2701	GGCACTCTGTATCTAGCTCAAGGTTTGTAAACACACCAACCAATCAGCACTCTGTATCTAGCT	2760
Db	2701	GGCACTCTGTATCTAGCTCAAGGTTTGTAAACACACCAACCAATCAGCACTCTGTATCTAGCT	2760
QY	2761	AGGATATGTGATATGACCAATTCAGCAGCTGTATCTGGTACTTTCAATGGGCATCGCTGT	2820
Db	2761	AGGATATGTGATATGACCAATTCAGCAGCTGTATCTGGTACTTTCAATGGGCATCGCTGT	2820

QY	2821	GAAGAGACCACCAACAGGCTTTGTGTGAGCAATTAAGCTTCTATCACCTGGGTGACGT	2880
Db	2821	GAAGAGACCACCAACAGGCTTTGTGTGAGCAATTAAGCTTCTATCACCTGGGTGACGT	2880
QY	2881	GGCGTACGCGAAAAAGAGAGTCACGCGCAAGGAGATTAAGGTGGGGCGCTTTATPAGAT	2940
Db	2881	GGCGTACGCGAAAAAGAGAGTCACGCGCAAGGAGATTAAGGTGGGGCGCTTTATPAGAT	2940
QY	2941	TTGGGTAGCTAAAGGAAAAATTACAGTCAAAGGGGGTTGTCTCTGGCGGCGACAGTGG	3000
Db	2941	TTGGGTAGCTAAAGGAAAAATTACAGTCAAAGGGGGTTGTCTCTGGCGGCGACAGTGG	3000
QY	3001	GGGGTCGCAAGGCTCTAGTGGGGGTCTTTTGTAGCCAGATGAGCCAGAAAAAGACT	3060
Db	3001	GGGGTCGCAAGGCTCTAGTGGGGGTCTTTTGTAGCCAGATGAGCCAGAAAAAGACT	3060
QY	3061	TTCCAAAGATATGTCAATTAAGGCAAGACCCGCAATTACCTCTTTGTGTG	3120
Db	3061	TTCCAAAGATATGTCAATTAAGGCAAGACCCGCAATTACCTCTTTGTGTG	3120
QY	3121	GAATGTCAATCAGTTAAGTTGGGGCGAGGCATATTCACCTCTTTGTGTGATTTTCAGTTAC	3180
Db	3121	GAATGTCAATCAGTTAAGTTGGGGCGAGGCATATTCACCTCTTTGTGTGATTTTCAGTTAC	3180
QY	3181	TTACAGGCCATCTGGCGGTATATGTGCCAAGTTACAGGGGATCGATGGCTTGGGCT	3240
Db	3181	TTACAGGCCATCTGGCGGTATATGTGCCAAGTTACAGGGGATCGATGGCTTGGGCT	3240
QY	3241	CAGAGGCTGACAGCTACTCTGTGGGGGCGCTTGGAGATGTTTGTGTGACACCTGAT	3300
Db	3241	CAGAGGCTGACAGCTACTCTGTGGGGGCGCTTGGAGATGTTTGTGTGACACCTGAT	3300
QY	3301	CTAGTTAATCTAGTGGGGAGCTGGAGAACCTTTTGTGTCTACTCAGGAGTTGTAAGCA	3360
Db	3301	CTAGTTAATCTAGTGGGGAGCTGGAGAACCTTTTGTGTCTACTCAGGAGTTGTAAGCA	3360
QY	3361	CCAACTACAGCGCTCTGTCAAAGACACCTCGGCTCTACCAATCAGCAGGATGTGGGTGG	3420
Db	3361	CCAACTACAGCGCTCTGTCAAAGACACCTCGGCTCTACCAATCAGCAGGATGTGGGTGG	3420
QY	3421	GGCCGATATAGGAATTAAGACAGGCTGCCGAGCGACATGTGGCAACGGCCACAGGCTC	3480
Db	3421	GGCCGATATAGGAATTAAGACAGGCTGCCGAGCGACATGTGGCAACGGCCACAGGCTC	3480
QY	3481	CTATCCAAATATGGCAGCTTGTCTTTTGTCTGTGGCGATTAATCTTGTCTACTGCTCG	3540
Db	3481	CTATCCAAATATGGCAGCTTGTCTTTTGTCTGTGGCGATTAATCTTGTCTACTGCTCG	3540
QY	3541	CTTTTGGGTCACACCTGCTTTTATAGCTGTAACTCACACGAAAGTCTGCAGCTTC	3600
Db	3541	CTTTTGGGTCACACCTGCTTTTATAGCTGTAACTCACACGAAAGTCTGCAGCTTC	3600
QY	3601	ACTCTGAAGCCACTTAAGACCAACGAGCCACCGGGAGATGAACAACTCCGGCGGCT	3660
Db	3601	ACTCTGAAGCCACTTAAGACCAACGAGCCACCGGGAGATGAACAACTCCGGCGGCT	3660
QY	3661	GCTTAAAGCTATTAACACTCACCGCGGAAGGTCTGTGACGCTTCACTCCTCAGCCAGCAGA	3720
Db	3661	GCTTAAAGCTATTAACACTCACCGCGGAAGGTCTGTGACGCTTCACTCCTCAGCCAGCAGA	3720
QY	3721	CCAGAACCCACAGAAAGAAATGCGAACAACATCTGAAACATCAGAAAGAACAAACT	3780
Db	3721	CCAGAACCCACAGAAAGAAATGCGAACAACATCTGAAACATCAGAAAGAACAAACT	3780
QY	3781	CCAGATGACACACTTAAGAGCTGTAACTCACTGGAGGGTCCGGGCTTCTTCTTG	3840
Db	3781	CCAGATGACACACTTAAGAGCTGTAACTCACTGGAGGGTCCGGGCTTCTTCTTG	3840
QY	3841	AAGTCAGAGACCAAGCACTCACAGTTTGGGACACAAAGCCGAGGATTTGAATAGC	3900
Db	3841	AAGTCAGAGACCAAGCACTCACAGTTTGGGACACAAAGCCGAGGATTTGAATAGC	3900
QY	3901	CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAATAAATTAACAAAATTTGGCGG	3960

|||||
Db 3901 CTGGGCAACATGTGAAATGCGCTCTGTGCAAAAAAAAAAAATTAACAAAATTGGCGG 3960
QY 3961 AGCATGTGTCCGTGCTGTGTGCTCCAGCTACGCGGAGGCTAAAGTGGGAGATGCT 4020
Db 3961 AGCATGTGTGCTGCTGTGTGCTCCAGCTACGCGGAGGCTAAAGTGGGAGATGCT 4020
QY 4021 TGAAGCTGGGAGGTGAAGACTGCATGAGTGTGATGTTACACAGCCCTTAGGCTGG 4080
Db 4021 TGAAGCTGGGAGGTGAAGACTGCATGAGTGTGATGTTACACAGCCCTTAGGCTGG 4080
QY 4081 GGACACACTGAGACCTGTTCCCTCCGCAAAAAAATTAACAAAAGTGAATAGAGT 4140
Db 4081 GGACACACTGAGACCTGTTCCCTCCGCAAAAAAATTAACAAAAGTGAATAGAGT 4140
QY 4141 GCGTGTATGTGCTAGGCGCAGTGGCTCATGCTGTATGCCACACTTTGGGAAAGCGAG 4200
Db 4141 GCGTGTATGTGCTAGGCGCAGTGGCTCATGCTGTATGCCACACTTTGGGAAAGCGAG 4200
QY 4201 GCGGGGCGGCTACCTTAAGTCAAGAGTGTGACACAGCCTGGCCAACTGGAAGAAAGCC 4260
Db 4201 GCGGGGCGGCTACCTTAAGTCAAGAGTGTGACACAGCCTGGCCAACTGGAAGAAAGCC 4260
QY 4261 ATCTCTTCAAAAATACAAAATTAAGCGGCTGTGGGGGCAAGTGGAGCATGCTGTAA 4320
Db 4261 ATCTCTTCAAAAATACAAAATTAAGCGGCTGTGGGGGCAAGTGGAGCATGCTGTAA 4320
QY 4321 TCCGACCTACTCAGAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGAGAGCGGCTTGC 4380
Db 4321 TCCGACCTACTCAGAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGAGAGCGGCTTGC 4380
QY 4381 AGTGACCCGAGATCGTGCCATTGCACTCCACCACTCCAGCCTGGGCAACAGCCAAA 4440
Db 4381 AGTGACCCGAGATCGTGCCATTGCACTCCACCACTCCAGCCTGGGCAACAGCCAAA 4440
QY 4441 CTCTGCTTAAAAAATAAAAAAATGCGCTGACATATTAAGAGGTGTGCAATGCAATAG 4500
Db 4441 CTCTGCTTAAAAAATAAAAAAATGCGCTGACATATTAAGAGGTGTGCAATGCAATAG 4500
QY 4501 TTGCGCAGGCAACATGTTTAAAGATGTGAGCTCTGCTCCCTTCAATGCTCTGTTAAAAAC 4560
Db 4501 TTGCGCAGGCAACATGTTTAAAGATGTGAGCTCTGCTCCCTTCAATGCTCTGTTAAAAAC 4560
QY 4561 CAGCCCAAGGCGGAGTGCATGCGCTCATGCTATATATCCACACTTTGGGAGCGCGAG 4620
Db 4561 CAGCCCAAGGCGGAGTGCATGCGCTCATGCTATATATCCACACTTTGGGAGCGCGAG 4620
QY 4621 GCGGGTGGATCACTGAGGTGAGAGTTCGAGACAGCCTGACCAACATGCTGAAT 4680
Db 4621 GCGGGTGGATCACTGAGGTGAGAGTTCGAGACAGCCTGACCAACATGCTGAAT 4680
QY 4681 CCCACCTCTAAAAAATACAAAATTAAGATGAGCATGTGTGCTGATGCCATCCAC 4740
Db 4681 CCCACCTCTAAAAAATACAAAATTAAGATGAGCATGTGTGCTGATGCCATCCAC 4740
QY 4741 CTACTTGGGAGGTGAGGCGAGGAAATCACTAGAACAGGAGCGGAGGTTTGTAGTGA 4800
Db 4741 CTACTTGGGAGGTGAGGCGAGGAAATCACTAGAACAGGAGCGGAGGTTTGTAGTGA 4800
QY 4801 CCGAGATGCTGCATGCACTCCAGCCTGAGCAATGAGGCAAACTCATCTCAAAAAAAC 4860
Db 4801 CCGAGATGCTGCATGCACTCCAGCCTGAGCAATGAGGCAAACTCATCTCAAAAAAAC 4860
QY 4861 AACAAACAAAACCACTCTACTCCAGGAGAGTGGTACAGAGCTGGGCCATCACT 4920
Db 4861 AACAAACAAAACCACTCTACTCCAGGAGAGTGGTACAGAGCTGGGCCATCACT 4920
QY 4921 GCAAGGTGCTGAGCCACAGAGCTAAAGCGGAGCTGAGAGCCCGGACCAAGATTAAGT 4980
Db 4921 GCAAGGTGCTGAGCCACAGAGCTAAAGCGGAGCTGAGAGCCCGGACCAAGATTAAGT 4980
QY 4981 TGTGAGATCAGTGTGAGATCAGAGCTGCTCCATGTTGATGACCAAGAGGGGCCCA 5040
Db 4981 TGTGAGATCAGTGTGAGATCAGAGCTGCTCCATGTTGATGACCAAGAGGGGCCCA 5040

Db 4981 TGTGAGATCAGTGTGAGATCAGAGCTGCTCCATGTTGATGACCAAGAGGGGCCCA 5040
QY 5041 AGCACCAGAGATGCCCCATCCAGTCAACACATCACTTCTCATCCAGAGATGCTGTT 5100
Db 5041 AGCACCAGAGATGCCCCATCCAGTCAACACATCACTTCTCATCCAGAGATGCTGTT 5100
QY 5101 CTTGGCAGCGTGGGGTAAATTAAGACAAAGGTGACAGTCTTGGGTGTGCTGAGTGA 5160
Db 5101 CTTGGCAGCGTGGGGTAAATTAAGACAAAGGTGACAGTCTTGGGTGTGCTGAGTGA 5160
QY 5161 TGCCCCAGGAGCGCTTGTGGCTGTAGAAAAGTTCAGAGCTTAGCCCGGACAGTGGC 5220
Db 5161 TGCCCCAGGAGCGCTTGTGGCTGTAGAAAAGTTCAGAGCTTAGCCCGGACAGTGGC 5220
QY 5221 TCAGCCCTGTAAATCCAGACACTTTGGGAGCGGAGCGGCTGTCAAGATCAGAGAGA 5280
Db 5221 TCAGCCCTGTAAATCCAGACACTTTGGGAGCGGAGCGGCTGTCAAGATCAGAGAGA 5280
QY 5281 TCGTGACCATCCCTGGCTAACACCGGTGAACCCCGTCTCTACTAAAAATNCAAAAATTGG 5340
Db 5281 TCGTGACCATCCCTGGCTAACACCGGTGAACCCCGTCTCTACTAAAAATNCAAAAATTGG 5340
QY 5341 CCGGGCATGCTGGCGGACCTGTAGTTCAGCTACTCGGAGGCTGAGGCAAGAGATG 5400
Db 5341 CCGGGCATGCTGGCGGACCTGTAGTTCAGCTACTCGGAGGCTGAGGCAAGAGATG 5400
QY 5401 GCGTGAACCCGAGAGCGAGATTTGCATGAGCGGAGATGCGCCACTGCATCCAGCT 5460
Db 5401 GCGTGAACCCGAGAGCGAGATTTGCATGAGCGGAGATGCGCCACTGCATCCAGCT 5460
QY 5461 GGGCGACAGACAACTCCATCTGGAAGAAAAGAAAAGAAACCTTCAGGTGACGAGA 5520
Db 5461 GGGCGACAGACAACTCCATCTGGAAGAAAAGAAAAGAAACCTTCAGGTGACGAGA 5520
QY 5521 GGGCCAGGCTGTATTTCTGTCACTTACATGACCTTGGGCAAGCACTTCTTCCGTGC 5580
Db 5521 GGGCCAGGCTGTATTTCTGTCACTTACATGACCTTGGGCAAGCACTTCTTCCGTGC 5580
QY 5581 CCAGTTCAGCGGCTTGAATGCACTCCAGAGCTCCCTTCAGATTAACCTGCATGGTTTC 5640
Db 5581 CCAGTTCAGCGGCTTGAATGCACTCCAGAGCTCCCTTCAGATTAACCTGCATGGTTTC 5640
QY 5641 TAAGATGAGAAATGGGGGCACTTCCCTCTCAACCCAGCCCGGTGTCACATCAAGT 5700
Db 5641 TAAGATGAGAAATGGGGGCACTTCCCTCTCAACCCAGCCCGGTGTCACATCAAGT 5700
QY 5701 GAATGACAGGAGAACTACGTTGCCAATCCCGCAGTTCCAAAGCCCTTGGGACCTTAC 5760
Db 5701 GAATGACAGGAGAACTACGTTGCCAATCCCGCAGTTCCAAAGCCCTTGGGACCTTAC 5760
QY 5761 TGTCAAGGCTGTCACAGAGAGGTGAAGGTCAAGGTGAGCCAAATCGCTCGAAGGCTTTC 5820
Db 5761 TGTCAAGGCTGTCACAGAGAGGTGAAGGTCAAGGTGAGCCAAATCGCTCGAAGGCTTTC 5820
QY 5821 CCTCATTTGGGAAACAGACATCCGTTTCTGTGCTCTACCGGGAATTCAGGGCTTTAGC 5880
Db 5821 CCTCATTTGGGAAACAGACATCCGTTTCTGTGCTCTACCGGGAATTCAGGGCTTTAGC 5880
QY 5881 CGAATGATCATGTGGGGGCGGGGGGTTTCTGGGGGAGTTCGCCAGCTTAATCACTTGGGA 5940
Db 5881 CGAATGATCATGTGGGGGCGGGGGGTTTCTGGGGGAGTTCGCCAGCTTAATCACTTGGGA 5940
QY 5941 CAGGACAGCCTGGAATCTTCAGATGTGCTCATCAAGTGTGGGTGGGCGACAGCCAA 6000
Db 5941 CAGGACAGCCTGGAATCTTCAGATGTGCTCATCAAGTGTGGGTGGGCGACAGCCAA 6000
QY 6001 GACCCAAATGCTTAACTCAGATGAGGCTCAGAGAGTTCGCCAGACAGGACCTTCGG 6060
Db 6001 GACCCAAATGCTTAACTCAGATGAGGCTCAGAGAGTTCGCCAGACAGGACCTTCGG 6060
QY 6061 AGAGTTTGGGGTAGAATGAGGAGCAACAGGCTTCTTTTCTCTTAAATTTGGG 6120
Db 6061 AGAGTTTGGGGTAGAATGAGGAGCAACAGGCTTCTTTTCTCTTAAATTTGGG 6120

Oy		6121	GGCTTGGGCGACAGGCTTTGACAATCCCAAAAGAGAGGGGCAAAAGACACTCCCCACAGAAG	6180
Dd		6121	GGCTTGGGCGACAGGCTTTGACAATCCCAAAAGAGAGGGGCAAAAGACACTCCCCACAGAAG	6180
Oy		6181	TCTGTCCGACAGCAGAGAGAGGAGACCCTCGACTGCACCTTCCCACAGAGCCT	6235
Dd		6181	TCTGTCCGAGAGGAAGAGAGGAGACCCTCGACTGCACCTTCCCACAGAGCCT	6235
<hr/>				
RESULT 2				
ID	AAZ29170			
XX	AAZ29170 standard; DNA: 6679 BP.			
XX	AAZ29170;			
XX	21-FEB-2000 (first entry)			
DE	Partial sequence of human G-CSF upstream of transcription start site.			
KW	Granulocyte colony stimulating factor; G-CSF; partial genomic sequence;			
KW	transcription start site; 5'non-coding sequence; DNA construct;			
KW	targeting sequence; regulatory region; marker gene; selection;			
KW	homologous recombination; gene therapy; delivery system; upstream;			
KV	haematopoietic progenitor cell; chemotherapy-induced neutropenia;			
KM	bone marrow transplantation; congenital neutropenic disorder;			
KX	chronic idiopathic disorder; ds.			
OS	Homo sapiens.			
FH	Key	Location/Qualifiers		
FT	misc_feature	1..19		
FT	/tag= a	/note= "Sequence derived from the junction of genomic		
FT	/insert and phage arm in the G-CSF/3 phage clone"	6550..6555		
FT	TATA_signal	/tag= b		
FT	CDS	/standard_name= "Hogness box"		
FT		6618..6679		
FT	/tag= c	/product= "N-terminal end of G-CSF"		
FT	/transl_except=	(Pos:6648, ACPATOSPMP)		
FT	/note=	"no stop codon given"		
PN	M09957291-A1.			
PD	11-NOV-1999.			
XX	05-MAY-1999;	99WO-USO9924.		
PR	07-MAY-1998;	98US-0084649.		
PA	(TRAN-) TRANSKARYOTIC THERAPIES INC.			
PI	Treco DA, Heartlein MW, Selden RF;			
DR	WPI: 2000-072235/06.			
DR	P-PsDB: AA15190.			
PT	Novel genomic sequences used for treating human diseases and disorders			
PS	-			
XX	Disclosure; Page 48-49; 58pp; English.			
CC	The present DNA sequence is the human partial genomic sequence from a			
CC	region upstream of the transcription start site of granulocyte colony-			
CC	stimulating factor (G-CSF). A DNA construct comprising a targeting			
CC	sequence homologous to the 5'/non-coding sequence of G-CSF, a			
CC	transcriptional regulatory sequence that differs from the endogenous			
CC	G-CSF gene, and a selectable marker gene for selection of recombinant			
CC	cells is generated, it can be used by homologous recombination to modify			
CC	the expression of endogenous G-CSF. These recombinant cells which express			
CC	G-CSF are useful for in vitro production of the protein and gene therapy.			

Such cells may also be used in a delivery system for stimulating the proliferation and differentiation of haematopoietic progenitor cells, or for other conditions that can be treated with G-CSF, like chemotherapy-induced neutropenia, to treat patients undergoing bone marrow transplantation, chronic idiopathic and congenital neutropenic disorders.

xx

SQ Sequence 6679 BP; 1684 A; 1739 C; 1830 G; 1426 T; 0 other:

Query Match	100.0%; Score 6235; DB 21; Length 6679;
Best Local Similarity	100.0%; Pred. NO. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY 1 GATCACTTGAGGACAGTACTTGAAGACCAAGCCTGGGACGATAGGACAGACTGCTCTACG 60	
DB	
DB 20 GATCACTTGAGGACAGTACTTGAAGACCAAGCCTGGGACGATAGGACAGACTGCTCTACG 79	
QY 61 AAAAAATCAAAAAATTAATGCGCGGCAATGGTGCTACGCTGTAAATCCCTGAACTTTGGG 120	
DB	
DB 80 AAAAAATCAAAAAATTAATGCGCGGCAATGGTGCTACGCTGTAAATCCCTGAACTTTGGG 139	
QY 121 ACATCAAGGCAAGTGGATCACTTTGAGAGTCAGAGATCGAGACTAGCCTGGCCAACTTGGT 180	
DB	
DB 140 ACATCAAGGCAAGTGGATCACTTTGAGAGTCAGAGATCGAGACTAGCCTGGCCAACTTGGT 199	
QY 181 GAACCCCTATCTCCACTTAAAAATACAAAAATTTAGCCAGCATGTGSCAGCAGCCTGTAA 240	
DB	
DB 200 GAACCCCTATCTCCACTTAAAAATACAAAAATTTAGCCAGCATGTGSCAGCAGCCTGTAA 259	
QY 241 ATCCGGGCTACTCAGAGGCTGAGCGAGGAGATCACTTGAACCCAGAGGCGGAGAGTTG 300	
DB	
DB 260 ATCCGGGCTACTCAGAGGCTGAGCGAGGAGATCACTTGAACCCAGAGGCGGAGAGTTG 319	
QY 301 CAGTAGCTGAATFACACACACCTGCACCTGCACCTGGGTGAGACAGACCAAGACTCTATCTC 360	
DB	
DB 320 CAGTAGCTGAATFACACACACCTGCACCTGCACCTGGGTGAGACAGACCAAGACTCTATCTC 379	
QY 361 AAAAAAAAAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420	
DB	
DB 380 AAAAAAAAAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 439	
QY 421 CTACTCAGGAGGCTGAGGTGGGAGATTCACCTTGAACCTGGGCGATCAAGGCTACAGTGA 480	
DB	
DB 440 CTACTCAGGAGGCTGAGGTGGGAGATTCACCTTGAACCTGGGCGATCAAGGCTACAGTGA 499	
QY 481 GCCAAGATCATCCCACTACACTCCAGCCCTGGGCAACAGAGAGAGACCTGTCTCTAAAAA 540	
DB	
DB 500 GCCAAGATCATCCCACTACACTCCAGCCCTGGGCAACAGAGAGAGACCTGTCTCTAAAAA 559	
QY 541 AATAATAATTAATTAAGAAAAAACAAGCTCTGTTATGTCTCTGCTGCTCATACTACTACT 600	
DB	
DB 560 AATAATAATTAATTAAGAAAAAACAAGCTCTGTTATGTCTCTGCTGCTCATACTACTACT 619	
QY 601 ATGTATATATGTTGCCAAACTCAAGATCCAGATAGTCAATTTTTATAGGCTTGTGGCCGT 660	
DB	
DB 620 ATGTATATATGTTGCCAAACTCAAGATCCAGATAGTCAATTTTTATAGGCTTGTGGCCGT 679	
QY 661 ATGGCTCTGTCACAAATCACTGCGCTGTCTTTAGACAAACAAACAGCACTATTAACAT 720	
DB	
DB 680 ATGGCTCTGTCACAAATCACTGCGCTGTCTTTAGACAAACAAACAGCACTATTAACAT 739	
QY 721 ACATACATGAATTTTTATAGACATCGAATTTGAATTTTCATATGATTTTTACATTTTAT 780	
DB	
DB 740 ACATACATGAATTTTTATAGACATCGAATTTGAATTTTCATATGATTTTTACATTTTAT 799	
QY 781 AAAAAATCTTTTTAAAAATTTTCCCTAACCATTTAAAAAGTGTAAAAACCGGCCAGCGC 840	
DB	
DB 800 AAAAAATCTTTTTAAAAATTTTCCCTAACCATTTAAAAAGTGTAAAAACCGGCCAGCGC 859	
QY 841 GCCATCGTCAGCGCCGTGAATTTCCAGACATTTTGGAGGCTGAGAGTGGGCGAGATCACTTGA 900	
DB	
DB 860 GCCATCGTCAGCGCCGTGAATTTCCAGACATTTTGGAGGCTGAGAGTGGGCGAGATCACTTGA 919	
QY 901 ATCAACAGTGTGAGGCGCCGACCATATGCAAAACCCGATTTCTCTATAAAATTTAA 960	

|||||
Db 920 ATCAACAGTTGAGACCGCTGGCCAAACATAGCAAAACCCCATTTCTACTAAAAATAAA 979
Qy 961 AAAATTTAGCTGGCATATAGTGTGCACACCTGTGATCCAGCTACTTGGAGGCTGAGCA 1020
Db 980 AAAATTTAGCTGGCATATAGTGTGCACACCTGTGATCCAGCTACTTGGAGGCTGAGCA 1039
Qy 1021 GGAGAAATCGCTTGAACCTGGGAGAGGGAGTTGGCAGTGAAGCAACATATGCTCACTGAC 1080
Db 1040 GGAGAAATCGCTTGAACCTGGGAGAGGGAGTTGGCAGTGAAGCAACATATGCTCACTGAC 1099
Qy 1081 TCACAGCTGGGTGACAGAGTGAACCTTGTCTCAACGAAAAAAAAGTGTAAAAGCCAT 1140
Db 1100 TCACAGCTGGGTGACAGAGTGAACCTTGTCTCAACGAAAAAAAAGTGTAAAAGCCAT 1159
Qy 1141 TCCAAATTCAGTGAACATCAATGTATACATACAGAGTGTGCTGCTGCTGAGGCAT 1200
Db 1160 TCCAAATTCAGTGAACATCAATGTATACATACAGAGTGTGCTGCTGCTGAGGCAT 1219
Qy 1201 ACCTGAGAAATGAGAGTTGCTTGTACAGAGCATACACATTTCCACATTAACTAGACAT 1260
Db 1220 ACCTGAGAAATGAGAGTTGCTTGTACAGAGCATACACATTTCCACATTAACTAGACAT 1279
Qy 1261 ACCAAGTTGCCATCCAAAGAGAGGTTTTTTTTTACAAATCTACATCCGCCAGCAACAAT 1320
Db 1280 ACCAAGTTGCCATCCAAAGAGAGGTTTTTTTTTACAAATCTACATCCGCCAGCAACAAT 1339
Qy 1321 GAGAGTTACTCCAGATCCTTTACAAAGATGCTTAAGCCCATACAGATATAAAACAGGA 1380
Db 1340 GAGAGTTACTCCAGATCCTTTACAAAGATGCTTAAGCCCATACAGATATAAAACAGGA 1399
Qy 1381 AGTGGAGAGGGAGACCTGCACAGCCCTTCTAAACATGAAGAAATACCTGTGAGAGCCCT 1440
Db 1400 AGTGGAGAGGGAGACCTGCACAGCCCTTCTAAACATGAAGAAATACCTGTGAGAGCCCT 1459
Qy 1441 GGATGCTGGAAGATGAATAACGGGGGTCTCTGAGAGCTGCCCTGTCAAGATCACTGTG 1500
Db 1460 GGATGCTGGAAGATGAATAACGGGGGTCTCTGAGAGCTGCCCTGTCAAGATCACTGTG 1519
Qy 1501 ACCTTGAAGCTCCAGTCCAGCTGTCAAGCCCATGTGTCAATGAGCATGATATGAGCCCT 1560
Db 1520 ACCTTGAAGCTCCAGTCCAGTCCAGCTGTCAAGCCCATGTGTCAATGAGCATGATATGAGCCCT 1579
Qy 1561 CACTCTCTGTTGGTCTTTATTCCTCCCATGTGGGGGTGAAGTGTGATTTGAGCCGTAT 1620
Db 1580 CACTCTCTGTTGGTCTTTATTCCTCCCATGTGGGGGTGAAGTGTGATTTGAGCCGTAT 1639
Qy 1621 TCAAGATGTACAGCTTTCTTGACAGGAAATAGTGTCAAGAAAAACAGAGGGGCTTGCA 1680
Db 1640 TCAAGATGTACAGCTTTCTTGACAGGAAATAGTGTCAAGAAAAACAGAGGGGCTTGCA 1699
Qy 1681 AGATGATCTAACTGCAAAATCCTTACCTGTGCTCAGGCACACCTAGTGTGTGATCTTGAC 1740
Db 1700 AGATGATCTAACTGCAAAATCCTTACCTGTGCTCAGGCACACCTAGTGTGTGATCTTGAC 1759
Qy 1741 AAGTTTTTCACTTCTCTGAGGCCATCCCTGGCTACACACACAGCAAGTTGGTTGACAGGA 1800
Db 1760 AAGTTTTTCACTTCTCTGAGGCCATCCCTGGCTACACACACAGCAAGTTGGTTGACAGGA 1819
Qy 1801 TGAATGAGAGATGCCCTTAACACTGTAACTCCAGCATTGGGAGGCCAAGGCGGGTGG 1860
Db 1820 TGAATGAGAGATGCCCTTAACACTGTAACTCCAGCATTGGGAGGCCAAGGCGGGTGG 1879
Qy 1861 ATGGGCTTGAAGCTGGAAGGTGACAGCATGCGGCACTCCACAGCCCTGCTGCTGCTC 1920
Db 1880 ATGGGCTTGAAGCTGGAAGGTGACAGCATGCGGCACTCCACAGCCCTGCTGCTGCTC 1939
Qy 1921 GGGGCTCTCTGCTGCTGGGGCTCCACTTGGTGGGCACTTGAAGAGCCCTTCAAGCCACCG 1980
Db 1940 GGGGCTCTCTGCTGCTGGGGCTCCACTTGGTGGGCACTTGAAGAGCCCTTCAAGCCACCG 1999
Qy 1981 CTGCACTGTGGAGGCCCTTTCTGAGGCTGGCAAGGCCAGAGCGGGCTCCCTAGCTTGC 2040
|||||

Db 2000 CTGCACTGTGGGAGGCCCTTTCTGGGCTTGGCCAGAGGCCAGCCGGCTCCTCAAGCTTGC 2059
Qy 2041 AGGAGGTGTGAGGAGAGGCTCAAGCAGGAACCGGGGCTTGGCAGCCGCTTGGCGGC 2100
Db 2060 AGGAGGTGTGAGGAGAGGCTCAAGCAGGAACCGGGGCTTGGCAGCCGCTTGGCGGC 2119
Qy 2101 CAGCTGGAGTTCGGGGTGGGCTTGGCGGGGCCCTGACACTGAGCAGCGGGCCAG 2160
Db 2120 CAGCTGGAGTTCGGGGTGGGCTTGGCGGGGCCCTGACACTGAGCAGCGGGCCAG 2179
Qy 2161 CCTGCGAGGCCCGGGCAATGAGAGGCTTAGCAACCGGGGCAAGGGGCTGGGAGGGTGT 2220
Db 2180 CCTGCGAGGCCCGGGCAATGAGAGGCTTAGCAACCGGGGCAAGGGGCTGGGAGGGTGT 2239
Qy 2221 ACTGGGTGCCCCAGCAGTGCAGCCCGCGGCTGTGCTGCTGATTTCTCACTGAGC 2280
Db 2240 ACTGGGTGCCCCAGCAGTGCAGCCCGCGGCTGTGCTGCTGATTTCTCACTGAGC 2299
Qy 2281 CTTAGCAGCCCTTCCCGGGGCAAGGCTTGGGAGCTGTGACGCCGCTAGTGCCTGCC 2340
Db 2300 CTTAGCAGCCCTTCCCGGGGCAAGGCTTGGGAGCTGTGACGCCGCTAGTGCCTGCC 2359
Qy 2341 CCTCATGGGCTCCTGTGTGGGCCCGAGGCTCCCGAGCAGCAGCCCTGCTCCACAG 2400
Db 2360 CCTCATGGGCTCCTGTGTGGGCCCGAGGCTCCTCCGAGCAGCAGCCCTGCTCCACAG 2419
Qy 2401 CGCCCATGCCATGACACAGCAGAAAGGCTGGAAGTGGGAGCGCAGCGCACCGGAGCTGG 2460
Db 2420 CGCCCATGCCATGACACAGCAGAAAGGCTGGAAGTGGGAGCGCAGCGCACCGGAGCTGG 2479
Qy 2461 CAGGACGCTACCCCTGGACCCCTGTGCGGAATCACTGGGTGAAGCAGCTGGGCTCCT 2520
Db 2480 CAGGACGCTACCCCTGGACCCCTGTGCGGAATCACTGGGTGAAGCAGCTGGGCTCCT 2539
Qy 2521 GAGCTGGTGAAGACTTGGAGAACTTTATGTCTAGCTCAGGGATCGAAATACACCAAT 2580
Db 2540 GAGCTGGTGAAGACTTGGAGAACTTTATGTCTAGCTCAGGGATCGAAATACACCAAT 2599
Qy 2581 CAGCACCCTGTGTCTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTGTATCTAGCT 2640
Db 2600 CAGCACCCTGTGTCTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTGTATCTAGCT 2659
Qy 2641 ACTGTGATGGGGCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATACACCAATC 2700
Db 2660 ACTGTGATGGGGCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATACACCAATC 2719
Qy 2701 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGACCCCTGTGTACTCTC 2760
Db 2720 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGACCCCTGTGTACTCTC 2779
Qy 2761 AGGATATGTAAATGACCAATTCGACAGTGTGTACTGTGCTACTTTCATGGGCAATCCGCT 2820
Db 2780 AGGATATGTAAATGACCAATTCGACAGTGTGTACTGTGCTACTTTCATGGGCAATCCGCT 2839
Qy 2821 GAAGAGACCAACCAAGGCTTGTGTGACCAATTAAGCTTCTATCACTGTGGGTGACAGT 2880
Db 2840 GAAGAGACCAACCAAGGCTTGTGTGACCAATTAAGCTTCTATCACTGTGGGTGACAGT 2899
Qy 2881 GGGCTGAGTCCGAAAAAGAGATGAGGAAAGGAGATTAAGGCTGGGGCCGTTTAATAGAT 2940
Db 2900 GGGCTGAGTCCGAAAAAGAGATGAGGAAAGGAGATTAAGGCTGGGGCCGTTTAATAGAT 2959
Qy 2941 TTGGGTATGTAAAGAAAAATTACAGTCAAAAGGGGTTTGTCTGTGGGGGCAAGAGTG 3000
Db 2960 TTGGGTATGTAAAGAAAAATTACAGTCAAAAGGGGTTTGTCTGTGGGGGCAAGAGTG 3019
Qy 3001 GGGGTGCAAGGTGCTCAGTGGGGGTCTTTTGAACCAAGATGACCAAGAAAGAGACT 3060
Db 3020 GGGGTGCAAGGTGCTCAGTGGGGGTCTTTTGAACCAAGATGACCAAGAAAGAGACT 3079
Qy 3061 TTCACAAGGTAAATGTATCAATTAAGGCAAGAACCCGCAATTTACACTCTTTTGTGTGG 3120
Db 3080 TTCACAAGGTAAATGTATCAATTAAGGCAAGAACCCGCAATTTACACTCTTTTGTGTGG 3139
|||||

QY	3121	GAATGTCATCAAGTTAAGTTGGGCGACGGCATATTCACCTCTTTTGTGATTTCTTGACGTTAC	3180
Db	3140	GAATGTCATCAAGTTAAGTTGGGCGACGGCATATTCACCTCTTTTGTGATTTCTTGACGTTAC	3199
QY	3181	TTACAGCCATCTGGGCGCTATATGTGCAAGTTACAGGGGATGCGATGGCTTGGGGCT	3240
Db	3200	TTACAGCCATCTGGGCGCTATATGTGCAAGTTACAGGGGATGCGATGGCTTGGGGCT	3259
QY	3241	CAGAGGCTTGCACGCTACTGTGGTGGGCGCTTGGAGAAATGTTGTGTGCACACTGTAT	3300
Db	3260	CAGAGGCTTGCACGCTACTGTGGTGGGCGCTTGGAGAAATGTTGTGTGCACACTGTAT	3319
QY	3301	CTAGTATATCTAGTGGGGACGTGGAGAACCTTTTGTCTTACTAGGATTTGTAAAGCA	3360
Db	3320	CTAGTATATCTAGTGGGGACGTGGAGAACCTTTTGTCTTACTAGGATTTGTAAAGCA	3379
QY	3361	CCAATCAGCGCCCTGTCAAACACAGACACTGGCTCAACATCAGCAGATGTGGGTG	3420
Db	3380	CCAATCAGCGCCCTGTCAAACACAGACACTGGCTCAACATCAGCAGATGTGGGTG	3439
QY	3421	GGCCAGATTAAGAGATTAAGACAGGCTGCCGAGCCAGCAGCTGCCAACGGCAGAGTCC	3480
Db	3440	GGCCAGATTAAGAGATTAAGACAGGCTGCCGAGCCAGCAGCTGCCAACGGCAGAGTCC	3499
QY	3481	CTATCCAAATATGGCAGCTTGTGTTCTTTTCTGTGGCGATTAATCTGTACTGTGCTG	3540
Db	3500	CTATCCAAATATGGCAGCTTGTGTTCTTTTCTGTGGCGATTAATCTGTACTGTGCTG	3559
QY	3541	CTTTTGGGCTCCACACTGCTTTTATAGCTGTAACTACACAGGAAGTCTGCAGCTTC	3600
Db	3560	CTTTTGGGCTCCACACTGCTTTTATAGCTGTAACTACACAGGAAGTCTGCAGCTTC	3619
QY	3601	ACTCCTGAAGCAGCTAAGACACGAGGCCACCGGAGGAATGAACTCCGGCGGCGCT	3660
Db	3620	ACTCCTGAAGCAGCTAAGACACGAGGCCACCGGAGGAATGAACTCCGGCGGCGCT	3679
QY	3661	GCCTTAAGAGCTATTAACACTCACCGGAAGTCTGCAGCTTCTCAGCCAGCCAGCA	3720
Db	3680	GCCTTAAGAGCTATTAACACTCACCGGAAGTCTGCAGCTTCTCAGCCAGCCAGCA	3739
QY	3721	CCAGGAACCCACGAGGAAGGAATCGGCAACAATCTGAACATCGAAGGAACAACCT	3780
Db	3740	CCAGGAACCCACGAGGAAGGAATCGGCAACAATCTGAACATCGAAGGAACAACCT	3799
QY	3781	CCAGATGACACCACTTAAGAGCTTAACTACATCGAGGAGTCCGGGCTTCTCTTGTG	3840
Db	3800	CCAGATGACACCACTTAAGAGCTTAACTACATCGAGGAGTCCGGGCTTCTCTTGTG	3859
QY	3841	AAGTCAGTGAACCAAGACACTCACCAAGTTTGGACACAGGCCAGAGTTTGAGATCAGC	3900
Db	3860	AAGTCAGTGAACCAAGACACTCACCAAGTTTGGACACAGGCCAGAGTTTGAGATCAGC	3919
QY	3901	CTGGGCAACATGATGAAGTGGCCCTCTGTGCAAAAAAATTTCAAAAAATTTGGCGG	3960
Db	3920	CTGGGCAACATGATGAAGTGGCCCTCTGTGCAAAAAAATTTCAAAAAATTTGGCGG	3979
QY	3961	AGCATGTGTGCTCGTGCCTGTGTGCCAGCTACCGGGAGGCTAAAGTGGGAGATCGCT	4020
Db	3980	AGCATGTGTGCTCGTGCCTGTGTGCCAGCTACCGGGAGGCTAAAGTGGGAGATCGCT	4039
QY	4021	TGAGCCGTGGGAGGTGAAGACTGCGATGATGTGATTTGTACACAGGCCCTCTAGGCTGGG	4080
Db	4040	TGAGCCGTGGGAGGTGAAGACTGCGATGATGTGATTTGTACACAGGCCCTCTAGGCTGGG	4099
QY	4081	GGACAGACTGAGACCCCTGTTTCCCTCCGCAAAAAAATTTACAAAAATGTAAATGAAGGT	4140
Db	4100	GGACAGACTGAGACCCCTGTTTCCCTCCGCAAAAAAATTTACAAAAATGTAAATGAAGGT	4159
QY	4141	GCCGTGATATGGCTTAGCGGCAAGTGGCTGATGCTGTAAATCCAGCACTTTTGGGAAGCCGAG	4200
Db	4160	GCCGTGATATGGCTTAGCGGCAAGTGGCTGATGCTGTAAATCCAGCACTTTTGGGAAGCCGAG	4219

QY	4201	GGGGGGGGGTACCTTAAGTACGAGAGTGTAGACCAAGCTGGCCACAATGAGAAAGCCC	4260
Db	4220	GGGGGGGGGTACCTTAAGTACGAGAGTGTAGACCAAGCTGGCCACAATGAGAAAGCCC	4279
QY	4261	ATCTCTCTTAAAAATACAAAATTAACCCGGCTGTGGGGGCAGTGGTGGAGCATGGCTGTAA	4320
Db	4280	ATCTCTCTTAAAAATACAAAATTAACCCGGCTGTGGGGGCAGTGGTGGAGCATGGCTGTAA	4339
QY	4321	TGCCAGACTCTCAGGAGGCTGAGGAGAGAACTACTTGAACCCAGAGAGCGCGGTTCG	4380
Db	4340	TGCCAGACTCTCAGGAGGCTGAGGAGAGAACTACTTGAACCCAGAGAGCGCGCGGTTCG	4399
QY	4381	AGTGAGCCGAGATCGTGCCATTGCATCTCACCCACTCCAGCTGGGCAAGAAGCCAAA	4440
Db	4400	AGTGAGCCGAGATCGTGCCATTGCATCTCACCCACTCCAGCTGGGCAAGAAGCCAAA	4459
QY	4441	CTCTCTCTTAAAAATAAAAATTAAGTGTCCAGATATTAAGAGGTGTGCATATG	4500
Db	4460	CTCTCTCTTAAAAATAAAAATTAAGTGTCCAGATATTAAGAGGTGTGCATATG	4519
QY	4501	TTCGACAGCAACATGTTTAAGAAATGTGAGACTCGCTTCATGCTGCTGTAAAAAC	4560
Db	4520	TTCGACAGCAACATGTTTAAGAAATGTGAGACTCGCTTCATGCTGCTGTAAAAAC	4579
QY	4561	CACCCCTCAAGGCCAGGTGCAGTGGCTATGCTTATATCCAGCATTTTGGAGGCCGAG	4620
Db	4580	CACCCCTCAAGGCCAGGTGCAGTGGCTATGCTTATATCCAGCATTTTGGAGGCCGAG	4639
QY	4621	GGGGGTGATCACTCGAGGTGAGGTTCGAGACCGCTGACCAACCACTGGTGAAT	4680
Db	4640	GGGGGTGATCACTCGAGGTGAGGTTCGAGACCGCTGACCAACCACTGGTGAAT	4699
QY	4681	CCCACTCTACTAAAAATACAAAATTAAGATGAGCATGGTGTGCATGGCTTAATCCAC	4740
Db	4700	CCCACTCTACTAAAAATACAAAATTAAGATGAGCATGGTGTGCATGGCTTAATCCAC	4759
QY	4741	CTACTTGGAGGCTGAGGCAGGAAATCACTAGAAACGAGGAGCGGAGTTGTAGTGAG	4800
Db	4760	CTACTTGGAGGCTGAGGCAGGAAATCACTAGAAACGAGGAGCGGAGTTGTAGTGAG	4819
QY	4801	CCGAAATGTGTGCATTTGCACTCCAGCCGACCAATGAGCCGAACCTCCATCCAAAAAAC	4860
Db	4820	CCGAAATGTGTGCATTTGCACTCCAGCCGACCAATGAGCCGAACCTCCATCCAAAAAAC	4879
QY	4861	AACAACAAAAAACCACCTCTACTCTCCAGGAGGCTGGGTAAAGAGCTGGGCCACATCAGT	4920
Db	4880	AACAACAAAAAACCACCTCTACTCTCCAGGAGGCTGGGTAAAGAGCTGGGCCACATCAGT	4939
QY	4921	GCAAGGTCTGTGACCAAGAGCTAAGGCGGAGCTGTGACAGACCAGGAGCAGATTAACGTG	4980
Db	4940	GCAAGGTCTGTGACCAAGAGCTAAGGCGGAGCTGTGACAGACCAGGAGCAGATTAACGTG	4999
QY	4981	TGTGAGATCAGTGTGTGAGATCAAGACGTCCCTGCCATTGGTGTACCAAGGGGCCCCCA	5040
Db	5000	TGTGAGATCAGTGTGTGAGATCAAGACGTCCCTGCCATTGGTGTACCAAGGGGCCCCCA	5059
QY	5041	AGCACAAGATAGGCCCCCATCCAGTCAACCAATCCACTTTCATCCAGAGATGTCTGTTT	5100
Db	5060	AGCACAAGATAGGCCCCCATCCAGTCAACCAATCCACTTTCATCCAGAGATGTCTGTTT	5119
QY	5101	CTTGGCAGCGTGGGGTAAATTAGACAGAAAGTGAAGTCTTGGGTGTGTCACTCAGAC	5160
Db	5120	CTTGGCAGCGTGGGGTAAATTAGACAGAAAGTGAAGTCTTGGGTGTGTCACTCAGAC	5179
QY	5161	TGCCCCAGGACGCGCTTGTGGCTGTAGAAAAGCTTAAGGCTTAAGGCCGGGCAACGCTGCG	5220
Db	5180	TGCCCCAGGACGCGCTTGTGGCTGTAGAAAAGCTTAAGGCTTAAGGCCGGGCAACGCTGCG	5239
QY	5221	TCAGGCTCTTAATCCACAGCACTTTTGGGAGGCGGAGGGGGGTATCAAGAGTCAAGAGA	5280
Db	5240	TCAGGCTCTTAATCCACAGCACTTTTGGGAGGCGGAGGGGGGTATCAAGAGTCAAGAGA	5299
QY	5281	TCGGTACCATCTGGCTTAACAAGGTGAAGCCCTCTCTCTACTAAAAATTAACAAAAATTTGG	5340

|||||
Db 5300 TCCTGACATCCCTGGCTAACACGCTGAAACCCCGTCTTACTAAATAACAAAAATTGG 5359
QY 5341 CCGGGATGGTGGCGGGACCTGTAGTTCAGCTACGCGGAGGCTGAGGACGAGAAATG 5400
Db 5360 CCGGGATGGTGGCGGGACCTGTAGTTCAGCTACGCGGAGGCTGAGGACGAGAAATG 5419
QY 5401 GCCTGAACCCGAGAGGAGAGATTGCAAGTGAAGCCGAGATGGCCCATGCACTCCAGCCT 5460
Db 5420 GCCTGAACCCGAGAGGAGAGATTGCAAGTGAAGCCGAGATGGCCCATGCACTCCAGCCT 5479
QY 5461 GGGCGACAGACGAAAGACTCTGGAAGAAAAAGAAACCTTACAGCTGAGCCAGA 5520
Db 5480 GGGCGACAGACGAAAGACTCTGGAAGAAAAAGAAACCTTACAGCTGAGCCAGA 5539
QY 5521 GGGCCAGGCTGTAATCTGTCACTTACATACATGACCTTGGGAGGCACTTCCCTGGC 5580
Db 5540 GGGCCAGGCTGTAATCTGTCACTTACATACATGACCTTGGGAGGCACTTCCCTGGC 5599
QY 5581 CCAGTTCACGGGGTTGGAATCGACTCCAGAGTCCCTTCCAGCATTAACGCTGATGGTTG 5640
Db 5600 CCAGTTCACGGGGTTGGAATCGACTCCAGAGTCCCTTCCAGCATTAACGCTGATGGTTG 5659
QY 5641 TAAGATGAGAAGATGGGGGCACTTCCCTCTCTACCCCAAGCCCGTGTCCACTTCAAGGT 5700
Db 5660 TAAGATGAGAAGATGGGGGCACTTCCCTCTCTACCCCAAGCCCGTGTCCACTTCAAGGT 5719
QY 5701 GAATGACACGAGGAATGACGTGTCCCAATCCCGAGTTCCAAAAGCCCTTGGGACCTTAC 5760
Db 5720 GAATGACACGAGGAATGACGTGTCCCAATCCCGAGTTCCAAAAGCCCTTGGGACCTTAC 5779
QY 5761 TGTGAGGCTGTGTCACGAGAGAGTGAAGTCAAGTGAAGCCATCGCTCGAAGGGCTTG 5820
Db 5780 TGTGAGGCTGTGTCACGAGAGAGTGAAGTCAAGTGAAGCCATCGCTCGAAGGGCTTG 5839
QY 5821 CCTCATTCGGGACGACATCGGTTTCTGCTGCTACCGGGATTCAGGGGCTTATAGC 5880
Db 5840 CCTCATTCGGGACGACATCGGTTTCTGCTGCTACCGGGATTCAGGGGCTTATAGC 5899
QY 5881 CGAATGATCATGGGGGCGGGGGGCTTTCGGGGGAGTTCCACGCTAATCACTTGGGA 5940
Db 5900 CGAATGATCATGGGGGCGGGGGGCTTTCGGGGGAGTTCCACGCTAATCACTTGGGA 5959
QY 5941 CAGGACAGCTGGAACCTTGCATGTGCTATCCAACTGTGGGGGCGACAGCCAA 6000
Db 5960 CAGGACAGCTGGAACCTTGCATGTGCTATCCAACTGTGGGGGCGACAGCCAA 6019
QY 6001 GACCAATGTCTTATCTCAAGTGTGGGGCTCAGAGGCTCCGACAGGCAAGCTCCGG 6060
Db 6020 GACCAATGTCTTATCTCAAGTGTGGGGCTCAGAGGCTCCGACAGGCAAGCTCCGG 6079
QY 6061 AGAGTTTGGGGGTAGGAATGGGAGCAACAGGCTTCTTTTCTCTTGAATTTGGG 6120
Db 6080 AGAGTTTGGGGGTAGGAATGGGAGCAACAGGCTTCTTTTCTCTTGAATTTGGG 6139
QY 6121 GCGTTGGGGGACAGGCTTGAAGATCCCAAAGGAGAGGGGCAAAAGGACACTCCCCACAG 6180
Db 6140 GCGTTGGGGGACAGGCTTGAAGATCCCAAAGGAGAGGGGCAAAAGGACACTCCCCACAG 6199
QY 6181 TCTGCCAGAGCGAGAGAGGAGAGACCCGACTCAGCTGCCACTTCCACAGGCT 6235
Db 6200 TCTGCCAGAGCGAGAGAGGAGAGACCCGACTCAGCTGCCACTTCCACAGGCT 6254

RESULT 3
AADI7443
ID AADI7443 standard; DNA: 6679 BP.

AC AADI7443;
XX

DT 10-DEC-2001 (first entry)
XX

DE Human granulocyte colony stimulating factor (hucsf) fragment #1.

XX Human: mutation; homologous recombination; target sequence: gene therapy;
KW homologous recombination-enhancing agent; non-homologous end joining;
KW therapeutic protein; granulocyte colony stimulating factor; hucsf; ds.
XX Homo sapiens.
XX WO200168882-A2.
XX 20-SEP-2001.
XX 13-MAR-2001; 2001MO-US07870.
XX 14-MAR-2000; 2000US-0525160.
XX (TRAN-) TRANSKARYOTIC THERAPIES INC.
XX Ivanov E;
XX WPI; 2001-582459/65.
XX
XX Complex or composition comprising a double stranded DNA sequence, a
PT homologous recombination-enhancing agent, and agent inhibiting
PT non-homologous end joining, for promoting alteration of a target
PT sequence in a cell -
XX
XX Disclosure; Page 77-79; 82pp; English.
XX
XX The invention relates to a complex for promoting alteration of a target
CC sequence in a cell, comprising a double stranded DNA sequence, a
CC homologous recombination-enhancing agent and an agent inhibiting non-
CC homologous end joining. The invention is used in gene therapy. The
CC complex is useful for promoting an alteration at a selected site of
CC a target sequence of a cell preferably of fungal, plant or animal
CC origin, or of vertebrate origin which is a primary or secondary
CC mammalian (human) cell or an immortalised mammalian (human) cell, where
CC target sequence comprises a mutation preferably point mutation having
CC less than 10 base pairs which differ from wild-type sequence, (selected
CC from cystic fibrosis transmembrane regulator (CFTR) gene having mutation
CC changes in an amino acid encoded by codon 508, beta-globin gene having
CC mutation changes in an amino acid encoded by codon 6, Factor VIII gene
CC having mutation changes in an amino acid encoded by codon 2209 or 2229,
CC Factor IX gene, von Willebrand factor gene or xeroderma pigmentosa group
CC G gene); and the DNA sequence comprises a wild-type sequence which can
CC correct the mutation. The method further comprises introducing an agent
CC which inhibits a mismatch-repair protein (expression), which is from
CC Msh2, Msh6, Msh3, Mhl1 and PMS2, or is an anti-mismatch-repair protein
CC antibody covalently linked to the DNA sequence, or to Rad52 protein or
CC its fragment. The complex is useful for altering expression of a protein
CC coding sequence of a gene in a cell. The method comprises introducing
CC the complex into the cell, where the DNA sequence comprises a regulatory
CC sequence, maintaining the cell under conditions which permit alteration
CC of a targeted genomic sequence to produce a homologously recombinant
CC cell and maintaining the homologously recombinant cell under conditions
CC which permit expression of the protein coding sequence of the gene under
CC control of the regulatory sequence. Homologously recombinant cell is
CC useful as a vehicle or delivery system for therapeutic proteins, such as
CC enzymes, hormones, cytokines, antigens, antibodies, clotting factors,
CC anti-sense RNA, regulatory proteins, transcription proteins, receptors,
CC structural proteins, novel (non-optimised) proteins and nucleic acid
CC products and engineered DNA and for supplying a therapeutic protein,
CC including erythropoietin, calcitonin, growth hormone, insulin and
CC insulinotropin. The present sequence is human granulocyte colony
CC stimulating factor (hucsf) fragment used in the invention.
XX
XX Sequence 6679 BP; 1684 A; 1739 C; 1830 G; 1426 T; 0 other;
SQ
Query Match 100.0%; Score 6235; DB 22; Length 6679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCACTTGAGAGACAGTAGTTCAAGACACGCTGGGAGCATAGGAGACTGTCTACG 60
|||||

Dh 20 GATCACTTGAGGACAGTAGTTCAGAACCCAGCTGGGACGATAGGAGAGACTGTCTACG 79
Qy 61 AAAAAATCAAAAAATTAATGCGCGGCGATGTTGGCTCAGCTCTGTAAATCCCTGAACCTTTGGG 120
Dh 80 AAAAAATCAAAAAATTAATGCGCGGCGATGTTGGCTCAGCTCTGTAAATCCCTGAACCTTTGGG 139
Qy 121 ACATCAAGGCAAGTAGATGATCTGAGGTGAGAGTTCAGAGTACCTGAGCCGCAACATGTT 180
Dh 140 ACATCAAGGCAAGTAGATGATCTGAGGTGAGAGTTCAGAGTACCTGAGCCGCAACATGTT 199
Qy 181 GAAACCCCTATCTCCACTAAAAAATATCAAAAAATTAAGCCAGGATGTTGGCAGGACCTGTA 240
Dh 200 GAAACCCCTATCTCCACTAAAAAATATCAAAAAATTAAGCCAGGATGTTGGCAGGACCTGTA 259
Qy 241 ATCCCGGCTACTCAGAGAGGCTGAGGAGAGATCATTGAACCCAGAGAGCGGAGGTTG 300
Dh 260 ATCCCGGCTACTCAGAGAGGCTGAGGAGAGATCATTGAACCCAGAGAGCGGAGGTTG 319
Qy 301 CAGTGAAGCTGAGATCACACCACTGCACTCCAGCCCTGGGTGACAGAGCAAGACTATCTC 360
Dh 320 CAGTGAAGCTGAGATCACACCACTGCACTCCAGCCCTGGGTGACAGAGCAAGACTATCTC 379
Qy 361 AAAAAAATTAATAAAAAATTAATAAAAAATTAAGCCAGGATGTTGGCAGGACCTGTA 420
Dh 380 AAAAAAATTAATAAAAAATTAATAAAAAATTAAGCCAGGATGTTGGCAGGACCTGTA 439
Qy 421 CTACTCAGAGAGGCTGAGGAGGATGATCTGAACCTGGGAGCTCAAGGCTACAGTGA 480
Dh 440 CTACTCAGAGAGGCTGAGGAGGATGATCTGAACCTGGGAGCTCAAGGCTACAGTGA 499
Qy 481 GCCAAGATCATGCCACTACCTCCAGCCCTGGGCAACAGAGAGACCCCTGCTCTAAAAA 540
Dh 500 GCCAAGATCATGCCACTACCTCCAGCCCTGGGCAACAGAGAGACCCCTGCTCTAAAAA 559
Qy 541 AATAATAATAATAAAAAAAGAACAGCTCTGTTATATGCTCTGCTCCATCATACTACT 600
Dh 560 AATAATAATAATAAAAAAAGAACAGCTCTGTTATATGCTCTGCTCCATCATACTACT 619
Qy 601 ATGATATATGTTTGCAGAACTCAAAAGATCCAGATAGTCAATTTTAAAGCTTGGCCGT 660
Dh 620 ATGATATATGTTTGCAGAACTCAAAAGATCCAGATAGTCAATTTTAAAGCTTGGCCGT 679
Qy 661 ATGCTCTCTGCACAACTCTGCCCTGCTTCTTACACAAAAAGAGCTATAAACAAT 720
Dh 680 ATGCTCTCTGCACAACTCTGCCCTGCTTCTTACACAAAAAGAGCTATAAACAAT 739
Qy 721 ACATACATGAATTTTATATAGACATCGAGATTGAATTCATATGATTTTACATTTTAT 780
Dh 740 ACATACATGAATTTTATATAGACATCGAGATTGAATTCATATGATTTTACATTTTAT 799
Qy 781 AAAAAATATCTTTAAAAAATTTCCCTTAACCATTTTAAAAAGTAAAAAGCCGCGCAGCC 840
Dh 800 AAAAAATATCTTTAAAAAATTTCCCTTAACCATTTTAAAAAGTAAAAAGCCGCGCAGCC 859
Qy 841 GCCATCGTCAGGCTGTAAATTCAGACACTTGGGAGGTGAGTGGGAGATCACTTGA 900
Dh 860 GCCATCGTCAGGCTGTAAATTCAGACACTTGGGAGGTGAGTGGGAGATCACTTGA 919
Qy 901 ATCAACAGTTGAGACCAAGCTGGCCCAACATAGCAAAACCCCATTTCTATAAAAATTA 960
Dh 920 ATCAACAGTTGAGACCAAGCTGGCCCAACATAGCAAAACCCCATTTCTATAAAAATTA 979
Qy 961 AAAAAATGCTGGCATAGTGTGCACACCTGTGATCCAGTACTTGGGAGGCTGAGCA 1020
Dh 980 AAAAAATGCTGGCATAGTGTGCACACCTGTGATCCAGTACTTGGGAGGCTGAGCA 1039
Qy 1021 GGAGAAATGCTTGAACCTGGGAGAGGAGTTCAGAGTACCAACATATGCACTGAC 1080
Dh 1040 GGAGAAATGCTTGAACCTGGGAGAGGAGTTCAGAGTACCAACATATGCACTGAC 1099
Qy 1081 TCCAGCCTGGGTGACAGAGTGAACCTGTCTCAACGAAAAAAGTGTAAAAAGCCAT 1140
Dh 1100 TCCAGCCTGGGTGACAGAGTGAACCTGTCTCAACGAAAAAAGTGTAAAAAGCCAT 1159

Qy 1141 TCCTAATTCAGTATCATCATGATCATACATCACTGAGGCTGCCGATCTGCTGTAGGAT 1200
Dh 1160 TCCTAATTCAGTATCATCATGATCATACATCACTGAGGCTGCCGATCTGCTGTAGGAT 1219
Qy 1201 ACCGTGAGAGTATGATGCTGTTGGTACAGAGCATCATATTTCCATTAATACATGACT 1260
Dh 1220 ACCGTGAGAGTATGATGCTGTTGGTACAGAGCATCATATTTCCATTAATACATGACT 1279
Qy 1261 ACCAAGTTCATCCATCAAGAGGTTTTTTTTTACAAATCTACATCCTCCCCAGCACAAT 1320
Dh 1280 ACCAAGTTCATCCATCAAGAGGTTTTTTTTTACAAATCTACATCCTCCCCAGCACAAT 1339
Qy 1321 GAGATCTACTCCAGATCCCTTTACAAAGATGCTCTAAGCCCACTACCAATGAAAAAG 1380
Dh 1340 GAGATCTACTCCAGATCCCTTTACAAAGATGCTCTAAGCCCACTACCAATGAAAAAG 1399
Qy 1381 AGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Dh 1400 AGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1459
Qy 1441 GGATCTGGAAGATGAATTAACGGGGGCTCTGAGAGCCTGCCCTGTCAATCACTGTG 1500
Dh 1460 GGATCTGGAAGATGAATTAACGGGGGCTCTGAGAGCCTGCCCTGTCAATCACTGTG 1519
Qy 1501 ACTTGTGAGCCTCCAGTCCAGTCTCAGCCCACTGTCTATGAGCCAGTATGAGCCCT 1560
Dh 1520 ACTTGTGAGCCTCCAGTCCAGTCTCAGCCCACTGTCTATGAGCCAGTATGAGCCCT 1579
Qy 1561 CACTCTGTTTGGTCTTATATCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Dh 1580 CACTCTGTTTGGTCTTATATCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1639
Qy 1621 TCAAGATGTACAGCTTTCTTGACAGAAAGTATGATCTCAGAAACAGCAGGAGGCTTGCA 1680
Dh 1640 TCAAGATGTACAGCTTTCTTGACAGAAAGTATGATCTCAGAAACAGCAGGAGGCTTGCA 1699
Qy 1681 AGATATCTACTGCAAAATCCACTGCTGACAGCCAGCAGCTATGTTCTGTAACTTTGAA 1740
Dh 1700 AGATATCTACTGCAAAATCCACTGCTGACAGCCAGCAGCTATGTTCTGTAACTTTGAA 1759
Qy 1741 AAGTTTTCACCTTCTGAGAGGCAATCCCTTGGTGTACAAACACAGGTTGTTGACAGA 1800
Dh 1760 AAGTTTTCACCTTCTGAGAGGCAATCCCTTGGTGTACAAACACAGGTTGTTGACAGA 1819
Qy 1801 TGAATGACGAAGTCCCTTACACCTGTAAATCCAGACACTTGGGAGGCCAAGCGGGTGG 1860
Dh 1820 TGAATGACGAAGTCCCTTACACCTGTAAATCCAGACACTTGGGAGGCCAAGCGGGTGG 1879
Qy 1861 ATGCTTGAAGCTGAGAGGTGACACATGCGGAGGTCTCTACACCCCTGCTGCTCTC 1920
Dh 1880 ATGCTTGAAGCTGAGAGGTGACACATGCGGAGGTCTCTACACCCCTGCTGCTCTC 1939
Qy 1921 GCGGCTCTCTGCTGAGGCTCCCACTTCGAGTGGAGCTTGAAGAGCCCTTACAGCCACG 1980
Dh 1940 GCGGCTCTCTGCTGAGGCTCCCACTTCGAGTGGAGCTTGAAGAGCCCTTACAGCCACG 1999
Qy 1981 CTGCACTGTGGAGCCCTTTCTGAGGCTGGCCAAAGGCAAGAGCCGCTCTCAGCTTGC 2040
Dh 2000 CTGCACTGTGGAGCCCTTTCTGAGGCTGGCCAAAGGCAAGAGCCGCTCTCAGCTTGC 2059
Qy 2041 AGGAGAGTGTGAGAGGAGAGGCTCAAGAGAGAACCGGGGCTGCGACGCGCTTCCGGGC 2100
Dh 2060 AGGAGAGTGTGAGAGGAGAGGCTCAAGAGAGAACCGGGGCTGCGACGCGCTTCCGGGC 2119
Qy 2101 CAGCTGAGATTCGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2160
Dh 2120 CAGCTGAGATTCGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2179
Qy 2161 CCTGTCAGGCCCCGGGCAATGAGAGGCTTACACCCGCGCCAGCGGCTGCGAGAGGTGT 2220
Dh 2180 CCTGTCAGGCCCCGGGCAATGAGAGGCTTACACCCGCGCCAGCGGCTGCGAGAGGTGT 2239

QY 2221 ACTGGTGGCCCCAGCAGTGGCCAGCCCGCGCGCTGTGCTGCTGATTTGCACTGGGC 2280
Db 2240 ACTGGGTGCCCCAGAGAGTGGCAGCCCGCGCGCTGTGCTGCTGATTTTCTCACTGGGC 2299
QY 2281 CTTAGCAGCCTTCCCGCGGGCAGGGGCTCGGGACCTGCAGCCCGCATGCTGAGCCCTCC 2340
Db 2300 CTTAGCAGCCTTCCCGCGGGCAGGGGCTCGGGACCTGCAGCCCGCATGCTGAGCCCTCC 2359
QY 2341 CCTCATGSGGCTGCTGTGCGGGCCGAGGCTCCCGAGCAGACCAACCCCTGTCTCAGAG 2400
Db 2360 CCTCATGSGGCTGCTGTGCGGGCCGAGGCTCCCGAGCAGACCAACCCCTGTCTCAGAG 2419
QY 2401 CGCCAGTCCCATGACACGACGCAAGGGCTGAGAAAGTGGGGCGCACGGCACCGGGAGCTGG 2460
Db 2420 CGCCAGTCCCATGACACGACGCAAGGGCTGAGAAAGTGGGGCGCACGGCACCGGGAGCTGG 2479
QY 2461 CAGGACGCTACCCCTGCAGCCCTGTGCGGAAATCCACTGSGTGAAGCCAGCTGGGCTCT 2520
Db 2480 CAGGACGCTACCCCTGCAGCCCTGTGCGGAAATCCACTGSGTGAAGCCAGCTGGGCTCT 2539
QY 2521 GAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCCGTAATATACCAAT 2580
Db 2540 GAGTCTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCCGTAATATACCAAT 2599
QY 2581 CAGCACCCTGTGTCTAGCTCAGGGTCTGTGAATGCACCAATCCACACTGTGTATCTAGCT 2640
Db 2600 CAGCACCCTGTGTCTAGCTCAGGGTCTGTGAATGCACCAATCCACACTGTGTATCTAGCT 2659
QY 2641 ACTGTGATGGGGCCCTTGGAGAACCTTTATGTCTAGCTCAGGGATTTGTAATACCAATC 2700
Db 2660 ACTGTGATGGGGCCCTTGGAGAACCTTTATGTCTAGCTCAGGGATTTGTAATACCAATC 2719
QY 2701 GGCACCTGTATCTAGCTCAGAGTTTGTAAACACACCAATCAGCACCCTGTGTAGCTC 2760
Db 2720 GGCACCTGTATCTAGCTCAGAGTTTGTAAACACACCAATCAGCACCCTGTGTAGCTC 2779
QY 2761 AGGCTATGTGAATCACCACATCGACAGTCTGTATCTGAGCTACCTTCATGGGCATCCGTGT 2820
Db 2780 AGGCTATGTGAATCACCACATCGACAGTCTGTATCTGAGCTACCTTCATGGGCATCCGTGT 2839
QY 2821 GAAGAGACCAACCAAGGGCTTTGTGAGCAATTAAGCCTTCTATCACCCTGGGTGCAGGT 2880
Db 2840 GAAGAGACCAACCAAGGGCTTTGTGAGCAATTAAGCCTTCTATCACCCTGGGTGCAGGT 2899
QY 2881 GGGCTGAGTCCGAAAAGAGATCAGCAGAGGAGATAGGGTGGGGCCGTTTATAGGAT 2940
Db 2900 GGGCTGAGTCCGAAAAGAGATCAGCAGAGGAGATAGGGTGGGGCCGTTTATAGGAT 2959
QY 2941 TTGGGTAGGTAAAGGAAAAATTACAGTCAAAAGGGGTTGTCTCTGCGGGCAGAGCTGG 3000
Db 2960 TTGGGTAGGTAAAGGAAAAATTACAGTCAAAAGGGGTTGTCTCTGCGGGCAGAGCTGG 3019
QY 3001 GGGGTGCAAGGTGCTAGTGGGGGTGCTTTTGAAGCAGAGATGAGCCAGAAAAAGACT 3060
Db 3020 GGGGTGCAAGGTGCTAGTGGGGGTGCTTTTGAAGCAGAGATGAGCCAGAAAAAGACT 3079
QY 3061 TTCAACAAGTAAATGTCACTAATTAAAGCAAGAACCCGCAATTACACCTCTTTTGTGTG 3120
Db 3080 TTCAACAAGTAAATGTCACTAATTAAAGCAAGAACCCGCAATTACACCTCTTTTGTGTG 3139
QY 3121 GAATGTCAATCAATTAAAGTTGGGCGAGGGCAATTACCTCTTTTGTGATTTCTTCACTTAC 3180
Db 3140 GAATGTCAATCAATTAAAGTTGGGCGAGGGCAATTACCTCTTTTGTGATTTCTTCACTTAC 3199
QY 3181 TTCAAGCCATCTGGGGGTATATGTGCAAGTTACAGGGGATGCATGGCTTGGGCTTGGCCT 3240
Db 3200 TTCAAGCCATCTGGGGGTATATGTGCAAGTTACAGGGGATGCATGGCTTGGGCTTGGCCT 3259
QY 3241 CAGAGGCTTGACAGCTACTCTGTGTGGGGCTTGGAGAAATGTTTGTGCGACACTGTGTAT 3300
Db 3260 CAGAGGCTTGACAGCTACTCTGTGTGGGGCTTGGAGAAATGTTTGTGCGACACTGTGTAT 3319
QY 3301 CTAGTTAAATCTAGTGGGAGCTGGAGAACCTTTGTGTAGCTCAGGGATTTGTAACGCA 3360

Db 3320 CTAGTTAAATCTAGTGGGAGCGTGGAGAACCTTTGTGTAGCTCAGGGATTTGTAACGCA 3379
QY 3361 CCAATCAGCGCCCTGTCAAAAACAGACACTGGGCTCTACCAATGACAGGATGTTGGGGG 3420
Db 3380 CCAATCAGCGCCCTGTCAAAAACAGACACTGGGCTCTACCAATGACAGGATGTTGGGGG 3439
QY 3421 GGCAGATTAAGAAATAAAGCAGGCTCCGAGCAGCAGTGGCAACGCCACAGCTCC 3480
Db 3440 GGCAGATTAAGAAATAAAGCAGGCTCCGAGCAGCAGTGGCAATGGCAACGCCACAGCTCC 3499
QY 3481 CTATCCAAATATGAGCAGCTTTGTCTTTTGTCTGTTCGATTAATCTTGTCTACTGCTG 3540
Db 3500 CTATCCAAATATGAGCAGCTTTGTCTTTTGTCTGTTCGATTAATCTTGTCTACTGCTG 3559
QY 3541 CTTTTGGGTCACACTGCTTTTATGAGCTTAACTACACACGAAAGGTCTGCAGCTC 3600
Db 3560 CTTTTGGGTCACACTGCTTTTATGAGCTTAACTACACACGAAAGGTCTGCAGCTC 3619
QY 3601 ACTCCTGAAGCCTATAGACACGAGCCACCGGAGAGATGAACAACTCCGGCCGCT 3660
Db 3620 ACTCCTGAAGCCTATAGACACGAGCCACCGGAGAGATGAACAACTCCGGCCGCT 3679
QY 3661 GCCTTAAGAGCTATTAACCTACCGCGAAGGCTGAGCTCAGCTCAGCCAGCCAGGA 3720
Db 3680 GCCTTAAGAGCTATTAACCTACCGCGAAGGCTGAGCTCAGCTCAGCCAGCCAGGA 3739
QY 3721 CCACGAACCCACGAAGGAAAGAACTGCGAACAATCTGAATCAGAAAGAACAACT 3780
Db 3740 CCACGAACCCACGAAGGAAAGAACTGCGAACAATCTGAATCAGAAAGAACAACT 3799
QY 3781 CCAATGACACCACTTAAAGGCTTAAACACTCAGAGGAGTCCGGGCTTCTTCTTG 3840
Db 3800 CCAATGACACCACTTAAAGGCTTAAACACTCAGAGGAGTCCGGGCTTCTTCTTG 3859
QY 3841 AAGTCAGTGAGACCAACCACTCAGCAGTTTGGGACACAGCCAGGATTTGAGATCAGC 3900
Db 3860 AAGTCAGTGAGACCAACCACTCAGCAGTTTGGGACACAGCCAGGATTTGAGATCAGC 3919
QY 3901 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTAACAAAATTTGGCGG 3960
Db 3920 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTAACAAAATTTGGCGG 3979
QY 3961 AGCATGGTGGTCCGTGCTGTGTCACAGTACCCGGGAGGCTTAAATGGAGATGCT 4020
Db 3980 AGCATGGTGGTCCGTGCTGTGTCACAGTACCCGGGAGGCTTAAATGGAGATGCT 4039
QY 4021 TGAACCTGGGAGGTGAAGACTGCACTGAGCTGTGATTTGTAACAGCCCTCTAGGCTGGG 4080
Db 4040 TGAACCTGGGAGGTGAAGACTGCACTGAGCTGTGATTTGTAACAGCCCTCTAGGCTGGG 4099
QY 4081 GGCAGACTGAGACCTGTTTCCCTCCGAAAAAAAATTGACAAAAGTGAATTAAGAGT 4140
Db 4100 GGCAGACTGAGACCTGTTTCCCTCCGAAAAAAAATTGACAAAAGTGAATTAAGAGT 4159
QY 4141 GCGTGAATATGAGTGAAGGCGAGTGGCTATGCTGTATCCACAGCCTTTGGGAAAGCCGAG 4200
Db 4160 GCGTGAATATGAGTGAAGGCGAGTGGCTATGCTGTATCCACAGCCTTTGGGAAAGCCGAG 4219
QY 4201 GCGGGCGGTACACTAAGGTCAAGAGTGTGAGACAGCAGCTGGCCCAACATGAGAAAGCCC 4260
Db 4220 GCGGGCGGTACACTAAGGTCAAGAGTGTGAGACAGCAGCTGGCCCAACATGAGAAAGCCC 4279
QY 4261 ATCTCTTCAAAAATAACAAAATTAGCCGGCTGTGGGGGAGTGTGGAGCATCCGTAA 4320
Db 4280 ATCTCTTCAAAAATAACAAAATTAGCCGGCTGTGGGGGAGTGTGGAGCATCCGTAA 4339
QY 4321 TCCACAGTACTGAGAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGAGGCGGGTTGC 4380
Db 4340 TCCACAGTACTGAGAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGAGGCGGGTTGC 4399
QY 4381 AGTGAGCCGAGATCGGCAATTGCACTCCACCCACTCAGCCCTGGGCAACAGAGCCAAA 4440

Db	4400	AGTGAGCCGAGATGCTGCATTGGCACTCCACCACCTCCAGGCTGGGCAACAAGGCCAA	4453
Qy	4441	CTGTGTCTTAAAAAAAAAAAAAAAAAGTGCCTGACATATTAAGAGGTGTGCATGCAATAG	4500
Db	4460	CTCTGTCTTAAAAAAAAAAAAAAAAAGTGCCTGACATATTAAGAGGTGTGCATGCAATAG	4519
Qy	4501	TTGGCAGGCACATGTTTAAACAATGTGAGAGCTCTGGCTTCCATGTGTCTTTAAAAAC	4566
Db	4520	TTGGCAGGCACATGTTTAAACAATGTGAGAGCTCTGGCTTCCATGTGTCTTTAAAAAC	4579
Qy	4561	CACCCCTCAAGGCGCAGAGTGCATGTGGCTTAAATCCAGACACTTTGGAGGCCGAG	4620
Db	4580	CACCCCTCAAGGCGCAGAGTGCATGTGGCTTAAATCCAGACACTTTGGAGGCCGAG	4639
Qy	4621	GGGGGTGATCACCTGAGTGCAGAGTGTTCGAGACCAGGCTACGCCACCAACATGTGTAAT	4680
Db	4640	GGGGGTGATCACCTGAGTGCAGAGTGTTCGAGACCAGGCTACGCCACCAACATGTGTAAT	4699
Qy	4681	CCCACTCTACTAAAAATACAAATTTAGATGACATGGTGGTGCATGCTGTAAATCCAC	4740
Db	4700	CCCACTCTACTAAAAATACAAATTTAGATGACATGGTGGTGCATGCTGTAAATCCAC	4759
Qy	4741	CTTACTTGGAGAGCTGAGGCGAGAAATACATAGAACCAAGGAGGGCGAGGTGTAGTAG	4800
Db	4760	CTTACTTGGAGAGCTGAGGCGAGAAATACATAGAACCAAGGAGGGCGAGGTGTAGTAG	4819
Qy	4801	CCGAGATCGTGCCACTTGTGCACCTGCAGCCTGAGCAATGAGCGAACTCCATCTCAAAAAAC	4866
Db	4820	CCGAGATCGTGCCACTTGTGCACCTGCAGCCTGAGCAATGAGCGAACTCCATCTCAAAAAAC	4879
Qy	4861	AACAACAAAAACCACCTCTCTAATCCAGGAGCTGGGTACAGAGCTGGGCCACATCAGT	4920
Db	4880	AACAACAAAAACCACCTCTCTAATCCAGGAGCTGGGTACAGAGCTGGGCCACATCAGT	4939
Qy	4921	GCAAGGTGCTAGGCGACAGAGCTAAAGGGGAGCTGCAGGACCAGGCGACCAATATTAACGTG	4980
Db	4940	GCAAGGTGCTAGGCGACAGAGCTAAAGGGGAGCTGCAGGACCAGGCGACCAATATTAACGTG	4999
Qy	4981	TGTGAGATCAGTGTGTGAGATCAGACGCTCCCTGCTAATGTGTGACACACAGGGGCCCA	5040
Db	5000	TGTGAGATCAGTGTGTGAGATCAGACGCTCCCTGCTAATGTGTGACACACAGGGGCCCA	5059
Qy	5041	AGCACAGAGATGGCCCCATCAGTCACACACATCCACTTCTCATCCAGAGATGTCTGTTT	5100
Db	5060	AGCACAGAGATGGCCCCATCAGTCACACACATCCACTTCTCATCCAGAGATGTCTGTTT	5119
Qy	5101	CTTGGGACGCTGGGGGTAAATTAGGACAGAGAGTACAGCTTGGGGTGTGTGCTCACTCAGAC	5160
Db	5120	CTTGGGACGCTGGGGGTAAATTAGGACAGAGAGTACAGCTTGGGGTGTGTGCTCACTCAGAC	5179
Qy	5161	TGCCCCAGGCGAGGCTGTGTGGCTGTAGAAACCTTAAGGCTTAAGGCCGAGGCGACAGTGGC	5220
Db	5180	TGCCCCAGGCGAGGCTGTGTGGCTGTAGAAACCTTAAGGCTTAAGGCCGAGGCGACAGTGGC	5239
Qy	5221	TCACGCTGTATCCACAGCACTTTGGAGGCGGAGGCGGGGTGATCAGAGAGTCAGAGGA	5280
Db	5240	TCACGCTGTATCCACAGCACTTTGGAGGCGGAGGCGGGGTGATCAGAGAGTCAGAGGA	5299
Qy	5281	TCGCGACATCTTGGGTAAACCGGTTGAACCCCGCTCTCTACTAAAAATACAAAAATTTGG	5340
Db	5300	TCGCGACATCTTGGGTAAACCGGTTGAACCCCGCTCTCTACTAAAAATACAAAAATTTGG	5359
Qy	5341	CCGGGCGATGTGTGGGGGACCTGTATGTTCCAGCTACTCCGAGAGCTGAGCGAGAGATG	5400
Db	5360	CCGGGCGATGTGTGGGGGACCTGTATGTTCCAGCTACTCCGAGAGCTGAGCGAGAGATG	5419
Qy	5401	GCGTGAACCCGAGAGCGAGAGTTTGCAGTGAACCGGAGATGCGGCCACTGCACCTCAGGCT	5460
Db	5420	GCGTGAACCCGAGAGCGAGAGTTTGCAGTGAACCGGAGATGCGGCCACTGCACCTCAGGCT	5479
Qy	5461	GGGGGACAGAGCAAGACTCCATCTCGGAAAAAAGAAAAACCTTCAGAGTGTGAGCCAGA	5520
Db	5480	GGGGGACAGAGCAAGACTCCATCTCGGAAAAAAGAAAAACCTTCAGAGTGTGAGCCAGA	5539

QY	5521	GGCCACAGGCTGTAATTCMTGCACTTACATACACTTGGGCAAGGCACTTCCTCCGCGC	5580
Db	5540	GGCCACAGGCTGTAATTCMTGCACTTACATACACTTGGGCAAGGCACTTCCTCCGCGC	5599
QY	5581	CCAGTTCACGGGGTGTGAATGCAGCTCCAAAGGTCCCTTCACACATTAACGCTGCATGGTTC	5640
Db	5600	CCAGTTCACGGGGTGTGAATGCAGCTCCAAAGGTCCCTTCACACATTAACGCTGCATGGTTC	5659
QY	5641	TAAATGTAAGAAAGATGGGGCAGTTTCCTCTCTCAACCCAGCCGCTGTCCACTTCAAGT	5700
Db	5660	TAAATGTAAGAAAGATGGGGCAGTTTCCTCTCTCAACCCAGCCGCTGTCCACTTCAAGT	5719
QY	5701	GAATACACAGGAAGATGCAGTGTCCCATATCCCGCAGTTCCCAAACCCCTTGGGGACCCATAC	5760
Db	5720	GAATACACAGGAAGATGCAGTGTCCCATATCCCGCAGTTCCCAAACCCCTTGGGGACCCATAC	5779
QY	5761	TGTCAGGGTCTGTCACAGAGAGAGTGAAGGTCAAGGTGAGCAGCAATCGCCTCGAAGGGTCTTG	5820
Db	5780	TGTCAGGGTCTGTCACAGAGAGAGTGAAGGTCAAGGTGAGCAGCAATCGCCTCGAAGGGTCTTG	5839
QY	5821	CCTCATTCGGGACAGACATCCGGTTTCCTCTGCTCTAACCAGGATTCAGAGGGCTTTAGC	5880
Db	5840	CCTCATTCGGGACAGACATCCGGTTTCCTCTGCTCTAACCAGGATTCAGAGGGCTTTAGC	5899
QY	5881	CGAATGATTCATGGGGGGGGGGGGGTTCTTGGGGGAGTTCGCCAGCTAATCACTTGGGA	5940
Db	5900	CGAATGATTCATGGGGGGGGGGGGGTTCTTGGGGGAGTTCGCCAGCTAATCACTTGGGA	5959
QY	5941	CAGACAGACCTTGAACTTTCATGGTGCCTATCCCAAGTGTGGGGGGGACAGACGCCAA	6000
Db	5960	CAGACAGACCTTGAACTTTCATGGTGCCTATCCCAAGTGTGGGGGGGACAGACGCCAA	6019
QY	6001	GACCCAAATGCTCTATCTCAGGTAGGGGCTCAGAGAAGTCTCCAGACAGGACGCTCCGG	6060
Db	6020	GACCCAAATGCTCTATCTCAGGTAGGGGCTCAGAGAAGTCTCCAGACAGGACGCTCCGG	6079
QY	6061	AGAGTTTGGGGGTAGGAATGGGAGAACCAAGGCTCTTTTTCCTTGAATTTGGG	6120
Db	6080	AGAGTTTGGGGGTAGGAATGGGAGAACCAAGGCTCTTTTTCCTTGAATTTGGG	6139
QY	6121	GGCTTTGGGGGACAGGCTTGAGAAATCCCAAGAGAGGGGCAAAAGGACACTCCCCACAAAG	6180
Db	6140	GGCTTTGGGGGACAGGCTTGAGAAATCCCAAGAGAGGGGCAAAAGGACACTCCCCACAAAG	6199
QY	6181	TCTGCGCAGAGCAGAGAGAGAGAGACCCGAGCTCAGCTGCACCTTCCCAAGGCTT	6235
Db	6200	TCTGCGCAGAGCAGAGAGAGAGAGACCCGAGCTCAGCTGCACCTTCCCAAGGCTT	6254
RESULT 4			
AADI17447			
ID	AADI17447	standard; DNA; 6753 BP.	
AC	AADI17447;		
XX	10-DEC-2001	(first entry)	
DE	Human interferon alpha2 (huIFNalpha2) fragment #4.		
KW	Human; mutation; homologous recombination; target sequence; gene therapy;		
KW	homologous recombination-enhancing agent; non-homologous end joining;		
KW	therapeutic protein; interferon alpha2; huIFNalpha2; ds.		
OS	Homo sapiens.		
XX	WO200168882-A2.		
PD	20-SEP-2001.		
PF	13-MAR-2001; 2001WO-US07870.		
PR	14-MAR-2000; 2000US-0525160.		

Dp	3058		TCGCTTTTGGGTCACACTGCTTTTAAAGAGCTTAACTACCTACACAGCAAGGTCTGCAGC	3117
Oy	3598		TTCACTCCTGTAAGCCACTTAAGACACAGCGCCACCGGGAGGAAATGAACAATCTCGGGCGC	3657
Dp	3118		TTCACTCCTGTAAGCCACTTAAGACACAGCGCCACCGGGAGGAAATGAACAATCTCGGGCGC	3177
Oy	3658		GCTCCCTTAAAGAGCTATTAACTACCTACCGCGGAAGGTCTGCAGCTTCACTCTCAGCCAGCG	3717
Dp	3178		GCTCCCTTAAAGAGCTATTAACTACCTACCGCGGAAGGTCTGCAGCTTCACTCTCAGCCAGCG	3237
Oy	3718		AGACCAGAACCCCAACGAGAGGAAGAACTGCGAACATTTGAAATCAGAGGAACAA	3777
Dp	3238		AGACCAGAACCCCAACGAGAGGAAGAACTGCGAACATTTGAAATCAGAGGAACAA	3297
Oy	3778		ACTCCAGATGCACACCTTAAAGACTGTAAACATCAGTCGAGAGGGTCCGGGCTCTCTCTC	3837
Dp	3298		ACTCCAGATGCACACCTTAAAGACTGTAAACATCAGTCGAGAGGGTCCGGGCTCTCTCTC	3357
Oy	3838		TTGAAGTCAGTGAGACCAAGCAGCTCACAGTTTGGACACAAAGCCAGAGATTGAGATC	3897
Dp	3358		TTGAAGTCAGTGAGACCAAGCAGCTCACAGTTTGGACACAAAGCCAGAGATTGAGATC	3417
Oy	3898		AGCCTGGGCAACATGATGAATGCTTCCCTCTCTGCAAAAAAAAAAAAAAAAAATTACAAAAATTGG	3957
Dp	3418		AGCCTGGGCAACATGATGAATGCTTCCCTCTCTGCAAAAAAAAAAAAAAAAAATTACAAAAATTGG	3477
Oy	3958		CGAGACATWGGGGTCCGTCGCTTGAGTCCCAAGCAGCGGGAGGCTAAAGTGGGAGATC	4017
Dp	3478		CGAGACATWGGGGTCCGTCGCTTGAGTCCCAAGCAGCGGGAGGCTAAAGTGGGAGATC	3537
Oy	4018		GCTTGAGCCTGGGAGGTGAAGACACTGCAAGTCACTGATTTGTACACAGCCCTCTAGACT	4077
Dp	3538		GCTTGAGCCTGGGAGGTGAAGACACTGCAAGTCACTGATTTGTACACAGCCCTCTAGACT	3597
Oy	4078		GGGGGACAGACTGAGACCTGTCTTCCCTCCGCAAAAAAAAAATTGCAAAAAGTGAATAGA	4137
Dp	3598		GGGGGACAGACTGAGACCTGTCTTCCCTCCGCAAAAAAAAAATTGCAAAAAGTGAATAGA	3657
Oy	4138		GGTCCCTGAAATGGCTTAGGGGCACTGGGCTCATGGCTTAATCCAGACACTTTGGGAAGCC	4197
Dp	3658		GGTCCCTGAAATGGCTTAGGGGCACTGGGCTCATGGCTTAATCCAGACACTTTGGGAAGCC	3717
Oy	4198		GAGGCGGGCGGGTCAAGGTCAAGGATCAGGAGTGTGAGACAGGCTGGCCAACTGAGAGAAG	4257
Dp	3718		GAGGCGGGCGGGTCAAGGTCAAGGATCAGGAGTGTGAGACAGGCTGGCCAACTGAGAGAAG	3777
Oy	4258		CCCATCTCTTTAAAAATACAAAATTAGCCGGCTGTGGGGGCACTGGTGGAGCATGCTG	4317
Dp	3778		CCCATCTCTTTAAAAATACAAAATTAGCCGGCTGTGGGGGCACTGGTGGAGCATGCTG	3837
Oy	4318		TAAATCCAGCTACTCAGAGGCTGTAGGCAAGAGATATACCTTGAACCCAGAGAGGGCGGCT	4377
Dp	3838		TAAATCCAGCTACTCAGAGGCTGTAGGCAAGAGATATACCTTGAACCCAGAGAGGGCGGCT	3897
Oy	4378		TGCAGTAGGCGGATGTGCTGATTTGCATCTCAACCACTCAGGCTGGGCAACAGAGCC	4437
Dp	3898		TGCAGTAGGCGGATGTGCTGATTTGCATCTCAACCACTCAGGCTGGGCAACAGAGCC	3957
Oy	4438		AAACTCTGTCTTAAAAAAAAAAAAAAAAAAGGCGCTGACATATTAAGAGGTGTGCATCAAA	4497
Dp	3958		AAACTCTGTCTTAAAAAAAAAAAAAAAAAAGGCGCTGACATATTAAGAGGTGTGCATCAAA	4017
Oy	4498		TAGTTGGCAGGCAACATGTTTAAGAATGTGAGGCTCTGCGCTTCATATGTTCTTTAAAA	4557
Dp	4018		TAGTTGGCAGGCAACATGTTTAAGAATGTGAGGCTCTGCGCTTCATATGTTCTTTAAAA	4077
Oy	4558		ACCCAGCCCTCAAGGCGCAGGTGACATGTGCTCATGCTTAATATCCAGACACTTTGGGAGGCC	4617
Dp	4078		ACCCAGCCCTCAAGGCGCAGGTGACATGTGCTCATGCTTAATATCCAGACACTTTGGGAGGCC	4137
Oy	4618		GAGGCGGGGTGATCACCTGAAGGTCAAGAGTTTCAGAGACAGCTTGACCAACATGTGTGA	4677

Db	4138	GAGCGGGGTGATCACTGAGGTGACGAGTTTCGAGACCAGCCCTGACCAACAATGGTGA	4137
OY	4678	AATCCACACCTCTACTAAATAATACAAATAATAGATGAGCATGGTGGTCATGCTGTAAATCC	4737
Db	4138	AATCCACACTCTACTAAATAATACAAATAATAGATGAGCATGGTGGTCATGCTGTAAATCC	4257
OY	4738	CACCTACTTGGGAGGCGTGAAGCAAAATCATTAGAACACAGGAGGCGGAGTTGTAGT	4797
Db	4258	CACCTACTTGGGAGGCGTGAAGCAAAATCATTAGAACACAGGAGGCGGAGTTGTAGT	4317
OY	4738	GAGCGGAAATCGTGGCATTTGCACTCCACACCTGAGCAATGAGCGAAATCTCCATTCAAAA	4857
Db	4318	GAGCGGAAATCGTGGCATTTGCACTCCACACCTGAGCAATGAGCGAAATCTCCATTCAAAA	4377
OY	4858	AACAACAACAAAAAACCACCTCTCTACTCCACAGGAGCGGGGTACAGACTGGGCGACATC	4917
Db	4378	AACAACAACAAAAAACCACCTCTCTACTCCACAGGAGCGGGGTACAGACTGGGCGACATC	4437
OY	4918	AGTGCMAAGTGTCTGAGCCACAGAGCTAAAGGCGAGCTTCAGAGACCGCGGACAGATTACA	4977
Db	4438	AGTGCMAAGTGTCTGAGCCACAGAGCTAAAGGCGAGCTTCAGAGACCGCGGACAGATTACA	4497
OY	4978	GTGTGTGAGATCAGTGTGTGAGATCAGACGTCCTGCCATTGGTGATCCACAGGGGGCCC	5037
Db	4498	GTGTGTGAGATCAGTGTGTGAGATCAGACGTCCTGCCATTGGTGATCCACAGGGGGCCC	4557
OY	5038	CCAAACACAGATGAGGCCCATCCAGCACACATCCACTTCATCCAGATGTCGTG	5097
Db	4558	CCAAACACAGATGAGGCCCATCCAGCACACATCCACTTCATCCAGATGTCGTG	4617
OY	5098	TTTTCTTGGCAGCGTGGGGTAAATTAGACAGAAAGTGAACAGTCTGGGTGTGTGACGTCA	5157
Db	4618	TTTTCTTGGCAGCGTGGGGTAAATTAGACAGAAAGTGAACAGTCTGGGTGTGTGACGTCA	4677
OY	5158	GACTGCCCCAGGACGAGCCTTGTGGCCTGTGAAAAAGCTTCAGGCGCTAGGCGGGCACGCT	5217
Db	4678	GACTGCCCCAGGACGAGCCTTGTGGCCTGTGAAAAAGCTTCAGGCGCTAGGCGGGCACGCT	4737
OY	5218	GGCTCAGCGCTGTAAATCCACACACTTTGGGAGCGGAGCGGGTGGATCAGAGGTCAAG	5277
Db	4738	GGCTCAGCGCTGTAAATCCACACACTTTGGGAGCGGAGCGGGTGGATCAGAGGTCAAG	4797
OY	5278	AGATCGTACCATCCTGGCTAAACAGAGGTGAACCCCGCTCTACTTAAATAATACAAAAAT	5337
Db	4798	AGATCGTACCATCCTGGCTAAACAGAGGTGAACCCCGCTCTACTTAAATAATACAAAAAT	4857
OY	5338	TGGCGGGGCATGTTGGCGGACCTGTATTCACAGTACTCGGAGGCGTGAAGCAGAGA	5397
Db	4858	TGGCGGGGCATGTTGGCGGACCTGTATTCACAGTACTCGGAGGCGTGAAGCAGAGA	4917
OY	5398	ATGGCGTGAACCCCGAGAGGCAAGTTTTGCAATGAGCCGAGATTCGCGCACTGCACTCCAG	5457
Db	4918	ATGGCGTGAACCCCGAGAGGCAAGTTTTGCAATGAGCCGAGATTCGCGCACTGCACTCCAG	4977
OY	5458	CCTGGGGCAGACAGACATCCATCTGTCGGAAGAAAAAGAAAAAGAAAAAGTTTCAGGTCGAGCC	5517
Db	4978	CCTGGGGCAGACAGACATCCATCTGTCGGAAGAAAAAGAAAAAGAAAAAGTTTCAGGTCGAGCC	5037
OY	5518	AGAGGCCAGGCTGTATTCTGTCACTTACATGACCTTGGCAAGGCACTTCCTCCCT	5577
Db	5038	AGAGGCCAGGCTGTATTCTGTCACTTACATGACCTTGGCAAGGCACTTCCTCCCT	5097
OY	5578	GGCCCAAGTTCAGGAGTGGGCAATGCACTCCAAGTCCCTTCACACATTTAAGCTGCAATGG	5637
Db	5098	GGCCCAAGTTCAGGAGTGGGCAATGCACTCCAAGTCCCTTCACACATTTAAGCTGCAATGG	5157
OY	5638	TTCTTAAGTGAAGATGAGGCAATTTCCCTCTCTCAACCCCAACCGGTGCACTTCATCA	5697
Db	5158	TTCTTAAGTGAAGATGAGGCAATTTCCCTCTCTCTCAACCCCAACCGGTGCACTTCATCA	5217
OY	5698	GGTGAATGACAGGGAAGTCAAGTGTCCCAATCCCGAGTTCCAAAGCCCTTGGGAGCC	5757
Db	5218	GGTGAATGACAGGGAAGTCAAGTGTCCCAATCCCGAGTTCCAAAGCCCTTGGGAGCC	5277

OY	5758	TACTGTCAAGGTCGTGCACACAGAGGGTAAGTAAAGTGAGCCCATGCCCTCGAAGGCTC	5817
Db	5278	TACTGTCTAGGGTCTGTGCACGAAGAAGGTGAAGGTGAGCCCAATGCCCCGTAAGGGCTC	5337
OY	5818	TTGGCTCATTTGGGGACAACATCCGGTTTTCTCGGCTCTAACCGGGAAATTCTAGGGGGCTT	5877
Db	5338	TTGGCTCATTTGGGGACAACATCCGGTTTTCTCTGGCTCTAACCGGGAAATTCTAGGGGGCTT	5397
OY	5878	AAGCCGAATGACTCATGGGGGGGGGGGGGGTTTCTGGGGGAGATTCCACGCTAATCAACTTG	5937
Db	5398	AGCCGAATGACTCATNNGGGGGGGGGGGGGTTTCTGGGGGAAATTCACGACTAATCAACTTG	5457
OY	5938	GGACAGCACAGCC	5950
Db	5458	GGATATTAGAGCC	5470

RESULT 5
AAZ29173

ID AA229173 standard; DNA; 2834 BP.

AC AAZ29173;

DT 21-FEB-2000 (first entry)

Targeting sequence-1 homologous to part of 5' non-coding region of G-CSF.

KM Granulocyte colony stimulating factor; G-CSF; genomic sequence; upstream;
KM transcription start site; 5' non-coding sequence; DNA construct;

KW homologous recombination; gene therapy; delivery system; CMV promoter; *hprt*

KW bone marrow transplantation; congenital neutropenic disorder;

2000

XX
XX
E000E7001-1

11 NOV 1966

XX
DE 05-MAY-1969 0010-12000004XX
07-MAY-1966XX
XX
XX (MAY -) MAY -)

XX
E

XX 2000-073333E 40C
DD

Novoj sekretar generalni
XX

xy
xy
xy

PS Disclosure; Fig 6; 58pp; English.

The present DNA sequence is the first targeting sequence that

CC a region upstream of the transcrip
CC stimulatory factor (C-CSE) a DNR

sequence is cloned into the plasmid pCG13, upstream of a CMV promoter and neomycin resistance gene that functions as the transcriptional start

regulatory sequence and as selectable marker respectively. Homologous recombination of this construct into the host cells can be used to

CC modify the expression of G-CSF. These recombinant cells which express
CC G-CSF are useful for *in vitro* production of the proteins and also the

such cells may also be used in a delivery system for stimulating the proliferation and differentiation of hematopoietic stem cells.

CC for other conditions that can be treated with G-CSF, like chemotherapy-induced neutropenia to protect patients undergoing bone marrow

transplantation, chronic idiopathic and congenital neutropenic disorders.

50	Sequence	2834 BP, 670 A; 759 C; 817 G; 588 T; 0 other:
	Query Match	45.5%; Score 2834; DB 21; Length 2834;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 2834;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1890	CCGGCAGTCTCACACAGCCCTGTTGCTCTGGCGCCTCCTCTGCTGGGCTCCCACTTC 1949
DB	1	CCGGCAGTCTCACAGCCCTGTTGCTCTGGCGCCTCCTCTGCTGGGCTCCCACTTC 60
QY	1950	GSTGGCACTTAGAGGAGCCCTTCAGGCGCACCGCTCAGCTGTTGGAGAGCCCTTTGCGGCTG 2009
DB	61	GGTGGCACTTAGAGGAGCCCTTCAGGCGCACCGCTCAGCTGTTGGAGAGCCCTTTGCGGCTG 120
QY	2010	GCCAAGGCGCAGAGCCGCGCTCCCTCAGCTTGGAGGAGGTGGAGGAGAGAGCTCAAGCA 2069
DB	121	GCCAAGGCGCAGAGCCGCGCTCCCTCAGCTTGGAGGAGGTGGAGGAGAGAGCTCAAGCA 180
QY	2070	GGAAACGGGGCTGCGCAGCGCGCTTGGCGGCGAGCTGGAGTTCCGGGTGGGCTGGGCTT 2129
DB	181	GGAAACGGGGCTGCGCAGCGCGCTTGGCGGCGAGCTGGAGTTCCGGGTGGGCGTGGCTT 240
QY	2130	GGCGGGGGCCCGCAGCTCGGAGAGAGGGGCGCAGACCCCTGCGAGGCCCCGGGGCAATGAGAGCT 2189
DB	241	GGCGGGGGCCCGCAGCTCGGAGAGAGGGGCGCAGACCCCTGCGAGGCCCCGGGGCAATGAGAGCT 300
QY	2190	TAGCACCCGGGGCCAGCGGCTCGGAGAGGAGGTACTGGGTGCGCCAGCAGTGCACACCGCC 2249
DB	301	TAGCACCCGGGGCCAGCGGCTCGGAGAGGAGGTACTGGGTGCGCCAGCAGTGCACACCGCC 360
QY	2250	GGCGCTGTGCTGCTCGATTTCCTCACTGGGCTTTAGCAGCCTTCCGCGGGGGCAGGCTC 2309
DB	361	GGCGCTGTGCTGCTGCTCGATTTCCTCACTGGGCTTTAGCAGCCTTCCGCGGGGGCAGGCTC 420
QY	2310	GGGACCTTCAGACCCGGCCATGCTGAGGCTGCCCTCCCTCCATGGGCTCTGTTGGGGCCGAGGC 2369
DB	421	GGGACCTTCAGACCCGGCCATGCTGAGGCTGCCCTCCCTCCATGGGCTCTGTTGGGGCCGAGGC 480
QY	2370	TCGCCGAGACACACCCCTCTGCTCCACAGACGGCCAGTCCCATCGACCAAGCAAGGGCT 2429
DB	481	TCGCCGAGACACACCCCTCTGCTCCACAGACGGCCAGTCCCATCGACCAAGCAAGGGCT 540
QY	2430	GAGAAATGCGGGCGCAGCGGACCGGAGCTGGCAGGCACTACCCCTCGAGCCTGTGGTCG 2489
DB	541	GAGAAATGCGGGCGCAGCGGACCGGAGCTGGCAGGCACTACCCCTCGAGCCTGTGGTCG 600
QY	2490	GAATCCAACTGGGGTAAAGCAGCTGGGGCTCGAGATCGTGGGAGACACTTGGAGAACCTTTA 2549
DB	601	GAATCCAACTGGGGTAAAGCAGCTGGGGCTCGAGATCGTGGGAGACACTTGGAGAACCTTTA 660
QY	2550	TGTCAGCTCAGGAGTCGTTAAATACCAATCAGCACACCTGTGCTACTCAGAGGCTGTGT 2609
DB	661	TGTCAGCTCAGGAGTCGTTAAATACCAATCAGCACACCTGTGCTACTCAGAGGCTGTGT 720
QY	2610	GAATGCACCAATCCACACTGTGTATCTAGCTACTCTGATGGGGCCCTTGGAGAACCTTTAT 2669
DB	721	GAATGCACCAATCCACACTGTGTATCTAGCTACTCTGATGGGGCCCTTGGAGAACCTTTAT 780
QY	2670	GTCATAGCTCAGGAGTGTAAATTAACAACAATGGGACACTGTGATATTAAGCTCAAGGTTTGTGA 2729
DB	781	GTCATAGCTCAGGAGTGTAAATTAACAACAATGGGACACTGTGATATTAAGCTCAAGGTTTGTGA 840
QY	2730	AACACACCAATCAGACACCTGTGTACTAGCTCAGGGTATGTGAATGACCAATGCACAGTC 2789
DB	841	AACACACCAATCAGACACCTGTGTACTAGCTCAGGGTATGTGAATGACCAATGCACAGTC 900
QY	2790	TGTATCTGGCTACTTTCTAATGGGCAATCCGCTGTGAAGAGACCAACCAAGGCTTTGTGTGA 2849
DB	901	TGTATCTGGCTACTTTCTAATGGGCAATCCGCTGTGAAGAGACCAACCAAGGCTTTGTGTGA 960
QY	2850	GCAATAAAGCTTCTATCACTCGGGTGCAGGTGGGCTAGTCCGAAAAAGAGTCAAGCGAA 2909
DB	961	GCAATAAAGCTTCTATCACTCGGGTGCAGGTGGGCTAGTCCGAAAAAGAGTCAAGCGAA 1020

QY	2910	GGGAGATTAAGGCTGGGGCCGTTTATATAGCATTTGGTAGTGAAGAAATTTACAGTCAA	2963
Db	1021	GGGAGATTAAGGCTGGGGCCGTTTATATAGGATTTGGTAGTGAAGAAATTTACAGTCAA	1080
QY	2970	AGGGGGTTTGTCTCTGGGGGGGAGAGTGGGGGGTGCAGAGGTCAGTGGGGGGTCT	3029
Db	1081	AGGGGGTTTGTCTCTGGGGGGGAGAGTGGGGGGTGCAGAGGTCAGTGGGGGGTCT	1140
QY	3030	TTTTGAGCCAGAGATGAGCCAGGAAAAGACTTTCACAAAGGTAAATGTCATCAATTAAGCA	3089
Db	1141	TTTTGAGCCAGAGATGAGCCAGGAAAAGACTTTCACAAAGGTAAATGTCATCAATTAAGCA	1200
QY	3090	AGGACCCGCCATTTTACACCTCTTTTGTTGGTGGAAATGTCATCAGTTAAGTGGGACAGGC	3149
Db	1201	AGGACCCGCCATTTTACACCTCTTTTGTTGGTGGAAATGTCATCAGTTAAGTGGGACAGGC	1260
QY	3150	ATATTCACTCTTTTGATTTCTGATCTCACTACTTTCAGGCCATCTGSGGGGTATATGTGCAAG	3209
Db	1261	ATATTCACTCTTTTGATTTCTGATCTCACTACTTTCAGGCCATCTGSGGGGTATATGTGCAAG	1320
QY	3210	TTTACAGGGGATGCATGGCTGGCTGTGGGCTCAGAGGCTTGACAGCTACTCTGTGGTGGGG	3269
Db	1321	TTTACAGGGGATGCATGGCTGGCTGTGGGCTCAGAGGCTTGACAGCTACTCTGTGGTGGGG	1380
QY	3270	CTTGAGAGAAATTTTGTGTGCACTCTGTATCTATGTTAAATCTAGTGGGGAGCTGAGAAC	3329
Db	1381	CTTGAGAGAAATTTTGTGTGCACTCTGTATCTATGTTAAATCTAGTGGGGAGCTGAGAAC	1440
QY	3330	CTTTGTGTAGCTCAGGGGATTTGTAAACGCACCAATACAGGCCCTGTCAAAACAGACCAC	3389
Db	1441	CTTTGTGTAGCTCAGGGGATTTGTAAACGCACCAATACAGGCCCTGTCAAAACAGACCAC	1500
QY	3390	TCGGCTCACCAAATCAGAGGATGTGGGTGGGGCAGATTAAGAAATPAAAGCAGGCTGC	3449
Db	1501	TCGGCTCACCAAATCAGAGGATGTGGGTGGGGCAGATTAAGAAATPAAAGCAGGCTGC	1560
QY	3450	CCGAGCCAGCAGTGGCAACGGCGCACAGGTCCCTATCCACAATATGGCAGCTTGTCTTT	3509
Db	1561	CCGAGCCAGCAGTGGCAACGGCGCACAGGTCCCTATCCACAATATGGCAGCTTGTCTTT	1620
QY	3510	TGCTGTTTGGCATTAATCTTGCTACTGTGCTTTTGGGTCCACACTGCTTTATGAGC	3569
Db	1621	TGCTGTTTGGCATTAATCTTGCTACTGTGCTTTTGGGTCCACACTGCTTTATGAGC	1680
QY	3570	TGTTAACTCTCACCGAAGGTCGTGACGTCCTACCTCCGTAAGCCACTAAGACACAGACCC	3629
Db	1681	TGTTAACTCTCACCGAAGGTCGTGACGTCCTACCTCCGTAAGCCACTAAGACACAGACCC	1740
QY	3630	ACCGGAGAGATGAAACAATCTCGGCGCGGCTTAAAGACTTAAACACTCAGCCGGAA	3689
Db	1741	ACCGGAGAGATGAAACAATCTCGGCGCGGCTTAAAGACTTAAACACTCAGCCGGAA	1800
QY	3690	GGTGTGACGCTTCACTCTCTCAGCCAGGAGACACGAAACCCACGAGGAAGAAACTGC	3749
Db	1801	GGTGTGACGCTTCACTCTCTCAGCCAGGAGACACGAAACCCACGAGGAAGAAACTGC	1860
QY	3750	GAAACAATCTTAACAATCAGAAAGAAACAATCTCCAGATGACACACTTAAAGAGCTGTAAACA	3809
Db	1861	GAAACAATCTTAACAATCAGAAAGAAACAATCTCCAGATGACACACTTAAAGAGCTGTAAACA	1920
QY	3810	CTCAGCTGGAGGGTCCGGCGCTTCTCTTAAAGTCAGTGGACCAAGACATCAACAGTT	3869
Db	1921	CTCAGCTGGAGGGTCCGGCGCTTCTCTTAAAGTCAGTGGACCAAGACATCAACAGTT	1980
QY	3870	TCGGACACAAAGCCAGAGATTGAGATCAGCTGGGCAACAATGATGAATATGCCCTCTCTG	3929
Db	1981	TCGGACACAAAGCCAGAGATTGAGATCAGCTGGGCAACAATGATGAATATGCCCTCTCTG	2040
QY	3930	CAAAAAAAAAAAATTTCAAAAATTTGGCGGACATGTTGGTCCGTGCTGTGGTCCAG	3989
Db	2041	CAAAAAAAAAAAATTTCAAAAATTTGGCGGACATGTTGGTCCGTGCTGTGGTCCAG	2100

OY		3990	CTAGCGGGAGGCTAAAGTGGAGATGCCCTGTAGCCCTGGAGGTGAAGAATCGACGAG	4004		
Db		2101	CTAGCGGGAGGCTAAAGTGGAGATGCCCTGTAGCCCTGGAGGTGAAGAATCGACGAG	2160		
OY		4050	CTGCATTGTACCACAGGCTCTTAGGCTGGGGGACAACACTGAGAACCTGTTTCCTCCCG	4109		
Db		2161	CTGTGATTGTACACAGGCTCTTAGGCTGGGGGACAACACTGAGAACCTGTTTCCTCCCG	2220		
OY		4110	CAAAAAAATTGACAAAAGTGTAAATAAGAGTGCCGTATATGCTTAGGCGCAGTGGCTCAT	4169		
Db		2221	CAAAAAAATTGACAAAAGTGTAAATAAGAGTGCCGTATATGCTTAGGCGCAGTGGCTCAT	2280		
OY		4170	GCCGTAAATCCCAGACACTTTGGGAAAGCCGAGAGCGGGGGTCACTTAAGGTCAAGAGTGT	4229		
Db		2281	GCCGTAAATCCCAGACACTTTGGGAAAGCCGAGAGCGGGGGTCACTTAAGGTCAAGAGTGT	2340		
OY		4230	GAGACACACCTGGGCCAATGATGAGAAAGCCCATCTCTTCTTAAATAATCAAAATTAAGCCGG	4289		
Db		2341	GAGACACACCTGGGCCAATGATGAGAAAGCCCATCTCTTCTTAAATAATCAAAATTAAGCCGG	2400		
OY		4290	CTGTGGGGGCGAGTGTGTGAGCATGCTCTTAATPCCAGCTACTCAAGAGGCTGAGGCAAGA	4349		
Db		2401	CTGTGGGGGCGAGTGTGTGAGCATGCTCTTAATPCCAGCTACTCAAGAGGCTGAGGCAAGA	2460		
OY		4350	GAATCACATTGAACCCAGAGAGCGGGGCTTGCACTGAGCCGAGATCGTCCACTTCC	4409		
Db		2461	GAATCACATTGAACCCAGAGAGCGGGGCTTGCACTGAGCCGAGATCGTCCACTTCC	2520		
OY		4410	AACCCATCCAGCTGGGGCAACAAGGCCAAACCTCGTCTTAAAAAATAAAAAAAAAAGTG	4469		
Db		2521	AACCCATCCAGCTGGGGCAACAAGGCCAAACCTCGTCTTAAAAAATAAAAAAAAAAGTG	2580		
OY		4470	CCTGCATATTAAGAGAGTGTGCATATAGTTGCCAGGCAACATGTTTTAAGATGTGGA	4529		
Db		2581	CCTGCATATTAAGAGAGTGTGCATATAGTTGCCAGGCAACATGTTTTAAGATGTGGA	2640		
OY		4530	GCTCTGCTCCATCATGTGCTGTAAAAAACCCACCCTCAAGGCCAGGTGCAGTGGCTCAT	4589		
Db		2641	GCTCTGCTCCATCATGTGCTGTGTAAAAAACCCACCCTCAAGGCCAGGTGCAGTGGCTCAT	2700		
OY		4590	GCCTTAATPCCAGACACTTTGGGAAAGCCGAGCGGGGTGGATCACTTAGGTCAAGAGTTC	4649		
Db		2701	GCCTTAATPCCAGACACTTTGGGAAAGCCGAGCGGGGTGGATCACTTAGGTCAAGAGTTC	2760		
OY		4650	GAGACACAGCTGACCAACCATGTGTAATCCCACTCTAATAAATAACAAATTAGA	4709		
Db		2761	GAGACACAGCTGACCAACCATGTGTAATCCCACTCTAATAAATAACAAATTAGA	2820		
OY		4710	TGAGCATGTGTGTG 4723			
Db		2821	TGAGCATGTGTGTG 2834			
<hr/>						
RESULT 6						
ID	AAD17444					
XX	AAD17444 standard; DNA; 2515 BP.					
AC	AAD17444;					
XX						
DT	10-DEC-2001	(first entry)				
XX						
DE	Human granulocyte colony stimulating factor (huGCSF) fragment #2.					
XX						
KM	Human; mutation; homologous recombination; target sequence; gene therapy;					
KW	homologous recombination-enhancing agent; non-homologous end joining;					
XX	therapeutic protein; granulocyte colony stimulating factor; huGCSF; ds.					
OS	Homo sapiens.					
PN	WO200168882-A2.					
XX						
PD	20-SEP-2001.					
XX						


```

Db 1261 ACCAAGTGGCCATCCAGAGAGGATTTTATTTTATCAATCTACACTCCCGCCAGCAACAAT 1320
QY GAGAGTTACTCCAGATCCCTTACAAAGATGCTCTAAGCCAGATACCAAGATGAACAGGA 1380
Db 1321 GAGAGTTACTCCAGATCCCTTACAAAGATGCTCTAAGCCAGATACCAAGATGAACAGGA 1380
QY 1381 AGTGGAGGGGAGAGCTGCGAGCCCTTCTAACCATGAAGAAATACCTGGTAGAGCCCTCT 1440
Db 1381 AGTGGAGGGGAGAGCTGCGAGCCCTTCTAACCATGAAGAAATACCTGGTAGAGCCCTCT 1440
QY 1441 GGATGCTCGAAGATGAATACAGGGGGTCTCTGAGAGCTCCCTGTCAGATCACTGTG 1500
Db 1441 GGATGCTCGAAGATGAATACAGGGGGTCTCTGAGAGCTCCCTGTCAGATCACTGTG 1500
QY 1501 ACTTGTAGCCTCAGAGCCAGTCCAGCCCATGTGTGACAGGCGAGATGAAGAGCCCT 1560
Db 1501 ACTTGTAGCCTCAGAGCCAGTCCAGCCCATGTGTGACAGGCGAGATGAAGAGCCCT 1560
QY 1561 CACTCTCTGTTGGTCTTTATTCCTCCCATGTGGGCTGAAGTCTGAGTTGAGCCCTTAT 1620
Db 1561 CACTCTCTGTTGGTCTTTATTCCTCCCATGTGGGCTGAAGTCTGAGTTGAGCCCTTAT 1620
QY 1621 TCAAGATGTACACTTTCTTGACAGGAAGTACTGTACAGAAAACAGAGGGCTTGGCA 1680
Db 1621 TCAAGATGTACACTTTCTTGACAGGAAGTACTGTACAGAAAACAGAGGGCTTGGCA 1680
QY 1681 AGATGATCTAACGCAAACTCTACCTGAGGCTCAGCCACCCACTGTTCTGTGATCTTAAAC 1740
Db 1681 AGATGATCTAACGCAAACTCTACCTGAGGCTCAGCCACCCACTGTTCTGTGATCTTAAAC 1740
QY 1741 AAGTTTTTCACTTCTCTGAGGCCATCCCTGGCTACACACACACAGATGGTTGACAGGA 1800
Db 1741 AAGTTTTTCACTTCTCTGAGGCCATCCCTGGCTACACACACACAGATGGTTGACAGGA 1800
QY 1801 TGAATAGACAGACTCCCTTACACCTGTATCCACAGCACTTTGGAGGCCAAGGGGGTGG 1860
Db 1801 TGAATAGACAGACTCCCTTACACCTGTATCCACAGCACTTTGGAGGCCAAGGGGGTGG 1860
QY 1861 ATGGCTTGAGCCGAGAGGTGACAGATGCGGCGACAGTCCACAGAGCCCTGCTGCTCTC 1920
Db 1861 ATGGCTTGAGCCGAGAGGTGACAGATGCGGCGACAGTCCACAGAGCCCTGCTGCTCTC 1920
QY 1921 GGGGCTCTCTGCTGAGGCTCCACTTGGTGGCACTTGGAGAGCCCTTCAAGCCACCG 1980
Db 1921 GGGGCTCTCTGCTGAGGCTCCACTTGGTGGCACTTGGTGGCACTTGGAGAGCCCTTCAAGCCACCG 1980
QY 1981 CTGCACTGTGGAAGCCCTTCTGAGGCTGCGCAAGGCCAGAGCCGGCTCCCTCAGCTTGC 2040
Db 1981 CTGCACTGTGGAAGCCCTTCTGAGGCTGCGCAAGGCCAGAGCCGGCTCCCTCAGCTTGC 2040
QY 2041 AGGGAGGTGTGGAGGGAGAGGCTCAAGACAGGAACCGGGGCTGGCGACGGGCTTGGCGGC 2100
Db 2041 AGGGAGGTGTGGAGGGAGAGGCTCAAGACAGGAACCGGGGCTGGCGACGGGCTTGGCGGC 2100
QY 2101 CAGCTGAGGTTCGGGGTGGGCGTGGGCTTGGGGGCGCCGCACTGCGAGAGAGGGGCGAG 2160
Db 2101 CAGCTGAGGTTCGGGGTGGGCGTGGGCTTGGGGGCGCCGCACTGCGAGAGAGGGGCGAG 2160
QY 2161 CCCTGCGAGGCCCGGGGCAATGAGAGGCTTAGACCCGGGCGCAGCGGCTCGGAGGGGTGT 2220
Db 2161 CCCTGCGAGGCCCGGGGCAATGAGAGGCTTAGACCCGGGCGCAGCGGCTCGGAGGGGTGT 2220
QY 2221 ACTGGGTGCC 2230
Db 2221 CAGGACAGCC 2230

```

```

RESULT 7
AA229174
ID AA229174 standard; DNA: 1252 BP.
XX
AC AA229174;
XX

```

```

DT 21-FEB-2000 (first entry)
XX
DE Targeting sequence-2 homologous to part of 5'non-coding region of G-CSF.
XX Granulocyte colony stimulating factor; G-CSF; genomic sequence;
XX translation start site; 5'non-coding sequence; DNA construct; downstream;
XX targeting sequence-2; regulatory region; marker gene; selection;
XX homologous recombination; gene therapy; delivery system; CMV promoter;
XX haematopoietic progenitor cell; chemotherapy-induced neutropenia;
XX bone marrow transplantation; congenital neutropenic disorder;
XX chronic idiopathic disorder; plasmid pGcl3; neomycin resistance gene; ds.
OS Homo sapiens.
XX
XX W09957291-A1.
XX
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-US09924.
XX
XX 07-MAY-1998; 98US-0084649.
XX
XX (TRAN-) TRANSKARYOTIC THERAPIES INC.
XX Treco DA, Heartlein MW, Selden RF;
XX
XX WPI: 2000-072235/06.
XX
XX Novel genomic sequences used for treating human diseases and disorders
XX
XX Disclosure; Fig 7; 58pp; English.
XX
XX The present DNA sequence is the second targeting sequence that
XX corresponds to nucleotides 4728-5979 of the human genomic sequence from
XX a region relative to the translation start site of granulocyte colony-
XX stimulating factor (G-CSF). A DNA construct comprising this targeting
XX sequence is cloned into the plasmid pGcl3, downstream of a CMV promoter
XX and neomycin resistance gene, that functions as the transcriptional
XX regulatory sequence and as selectable marker respectively. Homologous
XX recombination of this construct into the host cells, can be used to
XX modify the expression of G-CSF. These recombinant cells which express
XX G-CSF are useful for in vitro production of the protein and gene therapy.
XX Such cells may also be used in a delivery system for stimulating the
XX proliferation and differentiation of haematopoietic progenitor cells, or
XX for other conditions that can be treated with G-CSF, like chemotherapy-
XX induced neutropenia, to treat patients undergoing bone marrow
XX transplantation, chronic idiopathic and congenital neutropenic disorders.
XX
XX Sequence 1252 BP; 300 A; 337 C; 372 G; 243 T; 0 other:
XX
XX Query Match 20.1%; Score 1252; DB 21; Length 1252;
XX Best Local Similarity 100.0%; Pred. No. 5.5e-271;
XX Matches 1252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 4728 CTTGTAATCCCACTTCTTGGAGGCTGAGGCGAGAAATCTAGTAACAGGAGGCGG 4787
Db 1 CTTGTAATCCCACTTCTTGGAGGCTGAGGCGAGAAATCTAGTAACAGGAGGCGG 60
QY 4788 AGGTGTAGTGAAGCGGAGATCGTCCATTCGACTCCAGCCCTGAGCATGAGCGGAACTCC 4847
Db 61 AGGTGTAGTGAAGCGGAGATCGTCCATTCGACTCCAGCCCTGAGCATGAGCGGAACTCC 120
QY 4848 ATCTCAAAAAACAAACAACAAAAACCACTCTCTACTCCAGGAGGCTGGGTACAGAGCT 4907
Db 121 ATCTCAAAAAACAAACAACAAAAACCACTCTCTACTCCAGGAGGAGTGGGTACAGAGCT 180
QY 4908 GGGCCACATCAGTGCAGAGGTCTGAGCGACAGAGCTAAGCGGAGCTGCAGAGACCGCGGA 4967
Db 181 GGGCCACATCAGTGCAGAGGTCTGAGCGACAGAGCTAAGCGGAGCTGCAGAGACCGCGGA 240
QY 4968 CCAATTAACAGTGTGTGAGATCAAGTGTGTAGATCAAGACTCCCTGCATTTGGTACACAC 5027

```

Db 241 CCAGATTAACAGTGTGTGAGATCAGTGTGTGAGATCAGAGCTCCCTCCATTGGTGACAC 300
 QY 5028 CAGGGGGCCCCCAGCAGCAGAGATGCCCCCATTCAGTACAGATCCATTCTCATCCA 5087
 Db 301 CAGGGGGCCCCCAGCAGCAGAGATGCCCCCATTCAGTACAGATCCATTCTCATCCA 360
 QY 5088 GAGATGTCTGTTTCTTTGGGACCGCTGGGGTAAATTAAGACAGAAAGTGACAGTCTTGGGTG 5147
 Db 361 GAGATGTCTGTTTCTTTGGGACCGCTGGGGTAAATTAAGACAGAAAGTGACAGTCTTGGGTG 420
 QY 5148 TGTGAGTACAGATGAGTGGGACCGCTTGTGGCTGTAGAAAAGTTTCAGGCGCTAGGC 5207
 Db 421 TGTGAGTACAGATGAGTGGGACCGCTTGTGGCTGTAGAAAAGTTTCAGGCGCTAGGC 480
 QY 5208 CCGGACAGGTGTGCTACGCTCTGTAATCCAGCAGTCTTGGGAGCGCCGAGCGGTGATCA 5267
 Db 481 CCGGACAGGTGTGCTACGCTCTGTAATCCAGCAGTCTTGGGAGCGCCGAGCGGTGATCA 540
 QY 5268 CGAGGTACAGAGATGCTGACCATCTGCTGTAACACAGGTGAAACCCGCTCTACTTAAAA 5327
 Db 541 CGAGGTACAGAGATGCTGACCATCTGCTGTAACACAGGTGAAACCCGCTCTACTTAAAA 600
 QY 5328 TACAAAAAATTTGGGCGGCGATGCTGGCGGCGACCTGTACTTCCAGTACTCGGAGAGCTG 5387
 Db 601 TACAAAAAATTTGGGCGGCGATGCTGGCGGCGACCTGTACTTCCAGTACTCGGAGAGCTG 660
 QY 5388 AGCAGAGAGATGCGGTGAACCCGAGAGAGCAGAGTTCAGTGAAGCCGAGATCGCCAC 5447
 Db 661 AGCAGAGAGATGCGGTGAACCCGAGAGAGCAGAGTTCAGTGAAGCCGAGATCGCCAC 720
 QY 5448 TGCATCTCCAGCTGGGCGCAGAGCAGAACTCCATCTGSAAAAAGAAAAAGAAACGTCA 5507
 Db 721 TGCATCTCCAGCTGGGCGCAGAGCAGAACTCCATCTGSAAAAAGAAAAAGAAACGTCA 780
 QY 5508 GGTCTGAGCCAGAGGCGCCAGGCTTAATCTGTCATCTTCCATCCAGTCTTGGGAGAGCAC 5567
 Db 781 GGTCTGAGCCAGAGGCGCCAGGCTTAATCTGTCATCTTCCATCCAGTCTTGGGAGAGCAC 840
 QY 5568 TTCCTTCCTCGGCCAGTTCACGCGGGTTGAAATCGACTCCAAAGTCCCTTCAGCAGATTA 5627
 Db 841 TTCCTTCCTCGGCCAGTTCACGCGGGTTGAAATCGACTCCAAAGTCCCTTCAGCAGATTA 900
 QY 5628 CGCTGATGTGTTCTTAAGATGAGAAATGGGCGAGTTCCCTCTCTCACCCAGCCCGTG 5687
 Db 901 CGCTGATGTGTTCTTAAGATGAGAAATGGGCGAGTTCCCTCTCTCACCCAGCCCGTG 960
 QY 5688 TCCACTTCAAGGTGAATGACAGGGAAGTCACTGTGCCAATCCCGCACTTCCAAAGCCC 5747
 Db 961 TCCACTTCAAGGTGAATGACAGGGAAGTCACTGTGCCAATCCCGCACTTCCAAAGCCC 1020
 QY 5748 TTGGGGACCCCTACTGTCAAGGTCTGTCACAGAGAGGTGAAGGTCAAGTACGCCAATCGCC 5807
 Db 1021 TTGGGGACCCCTACTGTCAAGGTCTGTCACAGAGAGGTGAAGGTCAAGTACGCCAATCGCC 1080
 QY 5808 TCGAAGGGCTTGTGCTCATTTGGGAGACATCCGTTTCTCTGTGCTCTACCGGATTC 5867
 Db 1081 TCGAAGGGCTTGTGCTCATTTGGGAGACATCCGTTTCTCTGTGCTCTACCGGATTC 1140
 QY 5868 TGGGGCTTTTACCGCAATGATGAGTGGGCGGGGGGGTCTTCTGGGGAGTTCCAGCT 5927
 Db 1141 TGGGGCTTTTACCGCAATGATGAGTGGGCGGGGGGGTCTTCTGGGGAGTTCCAGCT 1200
 QY 5928 AATCAACTTGGGACAGGACAGCCTGGAATTTTCATGATGTGCTCATTCAGTGA 5979
 Db 1201 AATCAACTTGGGACAGGACAGCCTGGAATTTTCATGATGTGCTCATTCAGTGA 1252
 RESULT 8
 AAA96363/c
 ID AAA96363 standard; DNA; 50000 BP.
 XX
 AC AAA96363;
 XX

DT 08-FEB-2001 (first entry)
 XX
 DE Polymorphic repeat microsatellite sequences present in the CTLA4 locus.
 XX
 KW Autoimmune disease; polymorphic microsatellite repeat; PMR; CD28 gene;
 KW ICOS gene; CTLA4 gene; costimulatory receptor gene locus; CGRL; lupus;
 KW insulin-dependent diabetes mellitus; IDDM; Addison's disease; leprosy;
 KW Graves disease; autoimmune hypothyroidism; myasthenia gravis; thymoma;
 KW thyroiditis; postpartum thyroiditis; rheumatoid arthritis;
 KW Hashimoto's disease; coeliac disease; ss.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT satellite 5722..5746
 FT /tag= a
 FT /note= "sara41/42 microsatellite repeat"
 FT 6550..6597
 FT /tag= b
 FT /note= "sara43/44 microsatellite repeat"
 FT 19911..19956
 FT /tag= c
 FT /note= "PW210/211 microsatellite repeat"
 FT 23904..23957
 FT /tag= d
 FT /note= "sara45/46 microsatellite repeat"
 FT 27689..27780
 FT /tag= e
 FT /note= "sara17/18 microsatellite repeat"
 FT 30766..30801
 FT /tag= f
 FT /note= "sara19/20 microsatellite repeat"
 PN W0200056856-A2.
 PD
 XX 28-SEP-2000.
 XX
 PF 24-MAR-2000; 2000WO-0507938.
 XX
 PR 25-MAR-1999; 99US-0126215.
 XX
 PA (GENE) GENETICS INST INC.
 XX
 PI Ling V, Wu P, Gray GS;
 XX
 DR WPI; 2000-628257/60.
 XX
 PT Determining predisposition of humans to develop autoimmune disease
 PT Involves detecting polymorphic microsatellite repeat sequence within
 PT human costimulatory receptor gene locus -
 PS
 XX
 PS Claim 2; Page 67-82; 160pp; English.
 XX
 CC Two human bacterial artificial chromosome (BAC) clones that included
 CC and flanked the human CTLA4-4 locus were cloned and sequenced. The
 CC sequence data was assembled into a contiguous sequence that is presented
 CC in AAA96363-68. AAA96363-64 comprises BAC clone 22700, and AAA96365-68
 CC comprise BAC clone 22608. The sequences contain polymorphic
 CC microsatellite repeat (PMR) sequences. The specification describes a
 CC method for determining the predisposition of a human subject to develop
 CC autoimmune disease. The method comprises detecting a PMR sequence in the
 CC CD28, ICOS gene or CTLA4 gene of the human costimulatory receptor gene
 CC locus (hCGRL). PMR sequences vary in length among individuals and can be
 CC amplified to generate products that differ in size. These products can be
 CC then be detected by rapid and convenient high resolution processes. The
 CC method is useful for determining the predisposition of insulin-dependent
 CC diabetes mellitus (IDDM), Addison's disease, Graves disease, autoimmune
 CC hypothyroidism, myasthenia gravis, thymoma, lupus, thyroiditis,
 CC postpartum thyroiditis, rheumatoid arthritis, Hashimoto's disease,
 CC coeliac disease and leprosy. PMR sequences within hCGRL are useful as
 CC markers in a variety of assays and in the field of forensic medicine,
 CC disease diagnosis and human genome mapping.
 CC
 XX

Db 74236 |||||TGAGAGGTGACAGCTGTGCGACGTCTCAGAGCCCTGCTGTGCTCAGCACCTCTCT 74177
Qy 1933 GCCTGGGCTCCCACTTCCGTTGGTGGACCTTGGAGCCCTTACAGCCCAACCCGCTGACCTGTGG 1992
Db 74176 GCTTGGGCTCCCACTTGGGCGGCACTTGGAGAGCCCTTACAGCCCAACCTGACCTGTGG 74117
Qy 1993 AGCCCTTCTGTGGGCTGSCCAAGGCCAGGCCGCTCCCTCAGCTTGGAGGAGGTGG 2052
Db 74116 AGCCCTTCTGTGGGCTGSCCAAGGCCAGGCCACTCCCTCAGCTTGGAGGAGGTGG 74057
Qy 2053 AGGAGAGGCTCAACAGGAACCGGGGCTGGCGACGGGCTTGGCGGCGACGTGAGTTTC 2112
Db 74056 AGGAGAGGCGGAGCGGGAACCTGGGCTGCTCGAGGCTTGGCGGCGAGCTGAGTTTC 73997
Qy 2113 CGGGGGGCGTGGGCTTGGCGGAGCCCGCACTCGGAGCAGCGGGGAGCCCTGCGAGGCC 2172
Db 73996 CGGGGGGCGTGGGCTTGGCGGAGCCCGCACTCGGAGCAGCGGGGAGCCCTG -CTGGCC 73938
Qy 2173 CGGGGCAATGAGAGGCTTAGCACCCGGGCGAGCGGCTGCGAGGAGGTGACTGGTGGCC 2232
Db 73937 CGGGGCAATGAGAGGCTTAGCACCCGGGCGAGCGGCTGCGAGGAGGTGACTGGTGGCC 73878
Qy 2233 AGCAGTGGCAGCCCGCCCGCTGTGCTGCTCGATTTCTCAGTGGGCTTAGCAGCTT 2292
Db 73877 AGCAGTGGCAGCCCGCCCGCTGTGCTGCTCGATTTCTCAGTGGGCTTAGCAGCTT 73822
Qy 2293 CCCGGGGGAGGCGCTGGGACCTGACGCCGCACTGCTGAGCTT-----CCCTCCAT 2347
Db 73821 CCCGGGGGAGGCGCTGGGACCTGACGCCGCACTGCTGAGCTTCCACCCCTTCCAT 73762
Qy 2348 GGGCTCTGTGGGCGCCAGGCTCCCGAGACGACACACCCCTGCTCCACAGCCGCG 2407
Db 73761 GGGCTCTGTGGGCGCCAGGCTCCCGAGACGACACACCCCTGCTCCACAGCCGCG 73702
Qy 2408 TCCCATTCACCAACGAAAGGCTGAAAGTGGGGGCGACGGGACCGGGAAGTGGCGAGCG 2467
Db 73701 TCCCATTCACCAACGAAAGGCTGAAAGTGGGGAATGGCAAGCAGCGGCG -GAGAGCTGGCGGCGAG 73643
Qy 2468 CTACCCCTGACAGCCCTGCTGGGGAATTCACACTGGTGAAGCCAGTGGGGCTCTGAGTGG 2527
Db 73642 CTCACCTGACAGCCCTGCTGGGGAATTCACACTGGTGAAGCCAGTGGGGCTCTGAGTGG 73583
Qy 2528 GTGAGACTTGGAGAACCTTATGTCTAGCTCAGGGATCTTAATACACCAATCAGCAC 2587
Db 73582 GTGAGACTTGGAGAACCTTATGTCTAGCTCAGGGATCTTAATACACCAATCAGCAC 73523
Qy 2588 CTGTGTC-----TAGTCAAGGCTC 2606
Db 73522 CTGTGTCAGCTCAGGGTTTGTGAGTGCACCAATCAGCACCCCTGTGTAGCTCAGGGTT 73463
Qy 2607 TGTGAATGCACCAATTCACACTGTGTATCTAGCTACTGTGATGGGCTTGGAGAACCTT 2666
Db 73462 TGTGAATGCACCAATTCACACTGTGTATCTAGCTACTGTGATGGGCTTGGAGAACCTT 73403
Qy 2667 TATGCTTAGCTCAGGGATCTTAATACACCAATGGGCACTGTATCTAGCTCAAGTTT 2726
Db 73402 TATGCTTAGCTCAGGGATCTTAATACACCAATGGGCACTGTATCTAGCTCAAGTTT 73343
Qy 2727 GTAACACACCAATCAGACCCCTGTGTAGCTCAGGGATGTGTAATGACACCAATCAGCA 2786
Db 73342 GTAACACACCAATCAGACCCCTGTGTAGCTCAGGGATGTGTAATGACACCAATCAGCA 73283
Qy 2787 GTCTGTATCTGGCTACTTTCATGGGCACTCCGTGTGAAGAGACCAACCAAGGCTTGTG 2846
Db 73282 CTCTGTATCT----- 73273
Qy 2847 TGAGCAATTAAGCTTCTATCACTGGGTGCGAGTGGGCTGAGTCCGAAAAGAGACTCAGC 2906
Db 73272 ----- 73273
Qy 2907 GAAGGAGATTAAGGTGGGCGCTTTTATAGGATTTGGGTAGCTTAAGAAATTTACACT 2966

Db 73272 ----- 73273
Qy 2967 CAAGGGGGTTTGTCTCTGGCGGCGAGAGTGGGGGTCGCCAAGTGTCTAGTGGGGT 3026
Db 73272 ----- 73273
Qy 3027 GCTTTTGAAGCAGATGAGCCAGAAAGACTTTCACAAAGTAATGTCAATTAAG 3086
Db 73272 ----- 73273
Qy 3087 GCAGAGACCCGCAATTTACACTCTTTGTGGTGAATGTCAATGATTAAGTGGGGCAG 3146
Db 73272 ----- 73273
Qy 3147 GGCATATTCATCTTTTGTGATTTCTCAGTTACTTACAGGCCATCTGGGCTATATGTGC 3206
Db 73272 ----- 73273
Qy 3207 AAGTTACAGGGAGTCGATGCTTGGGCTGAGAGCTTGAACAGCTACTGTGGTGG 3266
Db 73272 -----AGATGCTGTGGTGG 73259
Qy 3267 GGCCTTGGAGAAATTTGTGTGACACCTGTATCTAGTTATCTAGTGGGAGCTGGAG 3326
Db 73258 GGCCTTGGAGAACTGTGTGAAAACCTGTATCTTACTAATCTGATGGGAGCTGGAG 73199
Qy 3327 AACCTTGTGTACTCAGGAGATTTGAACCCACCAATCAGCGCCCTGTCAAAACAGAC 3386
Db 73198 AACCTTGTGTACTCAGGAGATTTGAACCCACCAATCAGCGCCCTGTCAAAACAGAC 73139
Qy 3387 CACTGGGCTCTACCAATCAGCAGATGTGGTGGGCGAGATGAAGAAATTAAGCAGGC 3446
Db 73138 CACTGGGCTCTACCAATCAGCAGATGTGGGCGGCGAGATGAAGAAATTAAGCAGGC 73079
Qy 3447 TGCCTGGAGCAGCACTGGGAGCGGACAGGCTCCCTATTCACAAATATGGCAGCTTGTTC 3506
Db 73078 TGCCTGGAGCAGCACTGGGAGCGGCTCCCTGTCCACGCTGTGGAACTTGTTC 73019
Qy 3507 TTTTGTGTGTGCAATTAATCTTGTACTGCTCGGCTTTTGGGTCACACTCTTTTANG 3566
Db 73018 TTTTGTGTGTGCAATTAATCTTGTACTGCTCGCTCTTTTGGGTCACACTCTTTTANG 72959
Qy 3567 AGCTGTAACTCACCAGCAAGGTCTGCACTGCTTCACTCTGAAAGCCACTAAGACACGAG 3626
Db 72958 AGCTGTAACTCACCAGCAAGGTCTGCACTGCTTCACTCTGAAAGCCAGGAGACACGAG 72899
Qy 3627 CCCACCGGAGAAATGAACAACTCCGCGCGGCTGCTTAAGAGCTATTAACCTACGCG 3686
Db 72898 CCCACCGGAGGAGAAACAACTCCAGAGCGGCTGCTTAAGAGCTATTAACCTACGCG 72839
Qy 3687 GAAGGTCTGACGCTTCACTCTCAGCCAGCGAGACACGAAACCCACGAAAGAAAGAAAC 3746
Db 72838 GAAGGTCTGTAACCTCAGCTCGTAGACA--GAGACACGAAACCCACGAAAGAAAGAAAC 72781
Qy 3747 TEGGAAACATCTGAACATCAGAGAGAAACAACTCCAGATGACCACTTAAGACTGTA 3806
Db 72780 TCCGAAACATCTGAACATCAGAGAGAAACAACTCCAGAGCGGCCCACTTAAGACTGTA 72721
Qy 3807 ACACATCAGTGGAGGCTCCGCGGCTTCTGTGAAGTCAGTGAACCAAGCAGCACCA 3866
Db 72720 ACACATCAGTGGAGGCTCCAGGCTTCACTTCTTAAGTCAAGTGAACCAAGCAGCACCA 72661
Qy 3867 GTTTCGAGCACAAAGCCAGAGTTTGAATCAGCTGGGCAACATGATGAATGCGCTCT 3926
Db 72660 ATTTCGACATTAATAATCATCTTAAGAAATTAATTAAGTATTTGAATACTAAGCGCTGT 72601

RESULT 10
ABS56564/c

ID ABS56564 standard; DNA; 260209 BP.
XX
XX ABS56564;
XX
XX 27-JAN-2003 (first entry)
XX
XX Human SUFE2 genomic DNA sequence.
DE
XX SUFE2; ds; glucosamine-6-sulphatase; cancer; ischemia; human;
XX tumour; angiogenesis; coronary; carotid; arterial occlusive disease;
XX peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
XX thromboangitis obliterans; thrombotic disorder; vasculitis;
XX heart attack; myocardial infarction; vascular death; inflammation;
XX rheumatoid arthritis; asthma; adult respiratory distress syndrome;
XX sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
XX allograft rejection; lymphoma; thrombosis; sulphatase;
XX chromosome 20q12-13.2.
XX
XX Homo sapiens.
OS
XX
XX Key location/Qualifiers
XX misc_feature 184020..184561
XX /tag- a
XX /note- "This sequence represents 540
XX nucleotides not shown in the specification
XX as the result of a printing error"
XX
XX MO200259327-A2.
XX
XX 01-AUG-2002.
XX
XX 26-DEC-2001; 2001WO-US49793.
XX
XX 27-DEC-2000; 2000US-258577P.
XX 09-FEB-2001; 2001US-267831P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Rosen SD, Hemmerich S, Tomita M, Palmeri D;
XX WPI: 2002-636541/68.
XX
XX New sulfatase polypeptides having glucosamine-6-sulfatase activity,
XX useful in screening, discovery and preparation of diagnostic and
XX therapeutic agents for treating cancer, ischemic conditions,
XX inflammation, or thrombosis -
XX
XX Disclosure: Page 180-248; 293pp; English.
XX
XX This invention relates to the DNA and protein sequences of a novel
XX polypeptide having glucosamine-6-sulphatase activity. The sulphatases of
XX are useful in screening, discovery and preparation of diagnostic and
XX therapeutic agents for treating cancer, ischemic conditions,
XX inflammation, or thrombosis. The nucleic acids are useful in preparing
XX the sulfatase polypeptides, identifying the expression of genes in a
XX biological specimen, or generating transgenic non-human animals or
XX site-specific gene modification in cell lines. The host cells are
XX useful in replicating and/or expressing the polynucleotides or nucleic
XX acids. The agents are useful in treating the disorders cited above by
XX reducing tumour growth, inflammation, and thrombosis, or increasing
XX angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive
XX disease, peripheral arterial disease, atherosclerosis, myointimal
XX hyperplasia, thromboangitis obliterans, thrombotic disorders,
XX vasculitis; or preventing ischemic conditions, heart attack (myocardial
XX infarction), or other vascular death. The sulphatases and/or agents are
XX also useful in treating rheumatoid arthritis, asthma, adult respiratory
XX distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple
XX sclerosis, allograft rejection, and spread of lymphomas to cutaneous
XX sites. The present sequence represents the human SUFE2 genomic DNA
XX sequence of the invention. This gene maps to human chromosome
XX 20q12-13.2.
XX
XX Sequence 260209 BP; 78471 A; 51945 G; 51242 G; 77701 T; 850 other;

Query Match 11.9%; Score 744.8; DB 24; Length 260209;
Best Local Similarity 69.6%; Pred. No. 2.4e-156;
Matches 1399; Conservative 0; Mismatches 243; Indels 369; Gaps 14;
1875 AGAGGTGACAGCATCGCGGCACTGCTCCACAGCCCTGCTGCTGCGGCGCTCTGCG 1934
DB 222481 AGTCGTGACAGGCTGCTGCTGCTCCACAGCCCTGCTGCTGCGGCGCTCTGCG 222422
QY 1935 CTGGGCTCCCACTTGGTGGCACTTGGAGAGCCCTTCA-GCCACCGCTGCACCTGGGA 1993
DB 222421 GTAGGCTCCCACTTGGGCGGCACTTGGAGAGCCCTTCCGCGCGCTGCTGCGGA 222362
QY 1994 GCCCTTTTGGGCTGGGCAAGGCAAGAGCGGCTCCCTCAGCTTGCAGAGGGGTTGGA 2053
DB 222361 GCCCTTTTGGAGCTGGGCAAGGCAAGGCGGCTCCCTCAGCTTGCAGAGGGGTTGGA 222302
QY 2054 GGGAGAGGCTCAAGCAGAGAACCGGGGCTGCGACGCGGCTGCGGCGCAAGTGAAGTTCC 2113
DB 222301 GGGAGAGGCGGCAATGGGAAACCGGGGCTGCGGCTGCGGCGCAAGTGAAGTTCC 222242
QY 2114 GGGTGGGCGTGGGCTTGGGCGGCGCCGCACTGGAAGAGCGGCGCAAGCTTGCAGGCC 2173
DB 222241 GGATGGGCGTGGGCTTGGGCTGCGCCACACTGCGAGAGGCGCGCGGCTTGC-GGCC 222183
QY 2174 CGGGCAATGAGAGGCTTGAACCGCGGCGGCGGCGGCGGAGGTACTGGGTTGCCCA 2233
DB 222182 CGGGCAATGAGAGGCTTGAACCGCGGCGGCGGCGGCGGAGGTACTGGGTTGCCCA 222135
QY 2234 GCAGTGCACAGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2293
DB 222134 GCAGTGCACAGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 222079
QY 2294 CCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2348
DB 222078 CCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 222020
QY 2349 GGGTCTGCTGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2408
DB 222019 GGGTCTGCTGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221960
QY 2409 CCCATGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2468
DB 221959 CCCATGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 221901
QY 2469 TACCCCTGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2528
DB 221900 TCCACCTGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221841
QY 2529 TGGACACTTGGAGAACCTTATGCTAGCTCAGGAGATGTAATACCAATCAGCACGCC 2588
DB 221840 TGGGAGACGAGAG--TCTTATGCTAGCTCAGGAGATGTAATACCAATCAGCACGCC 221783
QY 2589 TGTGCTAGCTAGGAGCTGTAATGACCAATTCACACTGTAATCTAGTACTGTAT 2648
DB 221782 TGTGCTAGCTAGGAGCTGTAATGACCAATTCACACTGTAATCTAGTACTGTAT 221723
QY 2649 GGGGCTTGGAGAACCTTATGCTAGCTCAGGAGATGTAATACCAATCAGCACGCC 2708
DB 221722 GGGGCTTGGAGAACCTTATGCTAGCTCAGGAGATGTAATACCAATCAGCACGCC 221688
QY 2709 GTATCTAGCTCAAGTTTGTAAACACCAATCAGCACGCCCTGCTGCTAGCTAGGATG 2768
DB 221687 GTATCTAGCTCAAGTTTGTAAACACCAATCAGCACGCCCTGCTGCTAGCTAGGATG 221644
QY 2769 TGAATGACCAATGACAGCTGTAATGCTAGCTCAGTCTTATGAGGCAATGCGTGAAGAGC 2828
DB 221643 TGAATGACCAATGACAGCTGTAATGCTAGCTCAGTCTTATGAGGCAATGCGTGAAGAGC 221603
QY 2829 CACCAACAGGCTTGTGTAGACCAATTAAGCTTATACCTGCGGTGACAGGTGGCTGAG 2888
DB 221602 CACCAACAGGCTTGTGTAGACCAATTAAGCTTATACCTGCGGTGACAGGTGGCTGAG 221603

```

QY 2889 TCCGAAAAGAGTACAGGAGAGATAGGTTGGGCGCTTTTATAGATTTGGGTAG 2948
Db 221602 -----GCTTTGAGAGAC 221591
QY 2949 GTAAAGAAAATTACAGTCAAAAGGGGTTTGTCTGTGCGGAGAGAGTGGGGGTGCG 3008
Db 221590 CTTTATGTCTAGCTCAGGAGATTTGTAATACACATTTCAC----- 221551
QY 3009 AAGTGTCAAGTGGGGGTCTTTTGTAGCCAGATGAGCCAGAAAAGACTTTCACAG 3068
Db 221550 -----ACTGTATCTAG 221540
QY 3069 GTAAATGATCAATTAAGGAGACCGCCATTTACACTCTTTTGTGTGAGATGTCA 3128
Db 221539 CTCAAGGTTTGTAAACACACAAATCAGACCCCTGTGTACTCAAGTTTGGCAATGCA 221480
QY 3129 TCAATTAAAGTTGGGGCAGGGCATATTCATCTTTTGTGATTTCACTTACTTCAGGCC 3188
Db 221479 CCAATTCGA----- 221472
QY 3189 ATCTGGGCGTATATGTGCAAGTTACAGGGATGCGATGGCTTGGCTTCAGAGGCT 3248
Db 221471 -----CACTCTGT 221464
QY 3249 TGACAGCTACTGTGTGGGCGCTTGAGAAATGTTGTGTGACACTGTATCTAGTTAA 3308
Db 221463 ATCTAGCTGTCTGTGTGGGCGCTTGAGAAACTGTGTCTAAAATCTGTACTACTTAA 221404
QY 3309 TCTAGTGGGAGCTGAGAACTTTGTCTAGCTCAGAGGATTTTAAACGACCAATCAG 3368
Db 221403 TCTGATGGGAGAGTGGAACTTTGTATCTAGCTCAGAGGATTTTAAATGACCAATCAG 221344
QY 3369 CGGCGCTGCAAAAGACCACTGGCTCTACCATTCAGCAGAGATGTGGGTGGCCAGAT 3428
Db 221343 CACCTGTCAAAAGAGCCACTCGCTCTACCATTCAGCAGAGATGTGGGTGGCCAGAT 221284
QY 3429 AAGGATTAAGAGAGGCTGCGGAGAGAGATGAGGAGACCGGACAGACCTTCATCAC 3488
Db 221283 AAGGATTAAGAGAGGCTGCGGAGAGAGATGAGGAGACCGGAGACCTTCATCAC 221224
QY 3489 AATATGAGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3548
Db 221223 ACTGTGAGAGCTCGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 221164
QY 3549 GTCCACACCTCTTTTATAGCTGTAACTCAGCAGAGAGTCTGACGTTTACTCTTGA 3608
Db 221163 GTCCACGCGCTTTTATAGCTGTAACTCAGCAGAGAGATGTGTGTGTGTGTGTGTGT 221104
QY 3609 AGCCACTTAAGACACGAGCCGAGGAGATGAAATGAAATGAAATGAAATGAAATGAAAT 3668
Db 221103 GCCCAGGAGACACAGAGCCGAGGAGAGATGAAATGAAATGAAATGAAATGAAATGAAAT 221044
QY 3669 AGCTTAATACACTACCGGAGAGTGTGACGTTTACTCTCAGCAGCAGACAGACAGAAC 3728
Db 221043 AGCTTAATACACTCA-CGCGAAGGTGTGACGTTTACTCTCAGCAGCAGACAGACAGAAC 220985
QY 3729 CCACCGAAGGAAAGAACTCGGAACACATGTGAACATCAGAGGAAACAACTCCAGATGC 3788
Db 220984 CCACCGAAGGAAAGAACTCGGAACACATGTGAACATCAGAGGAAACAACTCCAGATGC 220925
QY 3789 ACCACTTAAGAGCTGTAACTACACTCAGAGGAGTCCGGGCTTCTTTTGAAGTACGT 3848
Db 220924 GCCACTTAAGAGCTGTAACTACACTCAGAGGAGTCCGGGCTTCTTTTGAAGTACGT 220865
QY 3849 GAGACCAAGCACTACCACTTTTGGACACAA 3879
Db 220864 GAGACCAAGCACTACCACTTTTGGACACAA 220834

```

RESULT 11
AB088143
ID AB088143 standard; cDNA; 121724 BP.

```

AC AB088143;
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 50.
XX
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
XX osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
PN WO200250301-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US48276.
XX
PR 18-DEC-2000; 2000US-255882P.
XX
PR 24-APR-2001; 2001US-285691P.
XX
XX
PA (GENE-) GENE LOGIC INC.
PA (PROC ) PROCTER & GAMBLE CO.
PI JI D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
DR WPI: 2002-557663/59.
XX
XX
XX Use of genes and their expression profiles associated with osteoblast
XX differentiation for screening modulators bone formation for diagnosing
XX or treating e.g. osteoporosis, or as markers for the differentiation
XX process.
XX
XX Claim 1: SEQ ID NO 50; 78bp + Sequence Listing; English.
XX
XX
XX The invention relates to genes and their expression profiles are used
XX for:
XX (a) screening modulators of precursor stem cell differentiation into
XX osteoblasts, or bone tissue deposition;
XX (b) diagnosing abnormal deposition of bone tissue, abnormal rate of
XX osteoblast formation or osteoporosis; or
XX (c) treating or monitoring treatment of the conditions cited in (b), or
XX monitoring the progression of bone tissue deposition.
XX Specific conditions include postmenopausal osteoporosis, glucocorticoid
XX osteoporosis or male osteoporosis, osteopenia, osteodystrophy,
XX drug-induced abnormalities in bone formation or bone loss, conditions
XX that involve altered bone metabolism (e.g. idiopathic juvenile
XX osteoporosis), skeletal disease linked to breast cancer, mastocytosis,
XX Fanconi syndrome or fibrous dysplasia. The present sequence is that of an
XX osteoblast differentiation associated cDNA marker of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 121724 BP; 33146 A; 25480 C; 26645 G; 36453 T; 0 other;
SQ

```

Query Match 11.8%; Score 733.2; DB 24; Length 121724;
Best Local Similarity 90.1%; Pred. No. 7.5e-154;
Matches 856; Conservative 0; Mismatches 78; Indels 16; Gaps 6;

```

QY 1867 TGACCTGAGAGGTGACAGATGCGGAGTCTCTACAGCCCTGTGCTTGGGCGC 1926
Db 4692 TGCACATGAGAGGTGAGACAGCTGTGCTTCCACAGCGCCGCTTCGCGCAC 4751
QY 1927 TCCCTGCGTGGGCTCCACTTGGTGGTGTGAGAGCCCTTTCAGCCACCGCTGCAC 1986
Db 4752 TCCCTGCGTGGGCTCCACTTGGTGTGACTTGAAGAGCCCTTTCAGCCACCGCTGCAC 4811
QY 1987 TGTGGAGCCCTTTTGTGGGCTGACCAAGGAGAGCGGCTCTCTACAGCTTGCAGAGAG 2046
Db 4812 TGTGGAGCCCTTTTGTGGGCTGACCAAGGAGAGCGGCTCTCTACAGCTTGCAGAGAG 4871
QY 2047 GTGTGAGAGGAGAGCTCAAGAGAAACGGGGCTGCGCAAGCGGCTTGGCGGCGACGCTG 2106

```

Db	4872	GTGTGGAGGAAGAGGCGCGCCCGGAACCGGGGGTGGCGCGGGCGCTTGTGGGGCAGCTG	49311
OY	2107	GAGTTCGGGGTGGGCGTGTGGGGGGCCCCGACATCGGAGCAGCGGGCAGCCCTGC	2166
Db	4932	GAGGTCCGGGGTGGGCGTGTGGGGGGCCCCGACATCGGAGCAGCGGGCGTCCCTGC	49911
OY	2167	CAGGCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGCAGCGGCTGGAGGGTGTACTGG	2226
Db	4992	C-GGCCCCGGGCAATGAGAGGCTTAGCACCCGGGCGCAGCACCTCGGAGAGTGTACTGG	5050
OY	2227	TGCCCCAGACAGTGGCAGCGCGCGGGCGGTGGTGGTGGCGATTTCTCACTGGGCGTTAGC	2286
Db	5051	TCCTCCACAGTGGCAGCGCGCGGGCGGTGGTGGTGGCGATTTCTCACTGGGCGTTAGC	5106
OY	2287	AGCCTTCGCCGGGGGAGGCGCTCGGGACCTGACAGCCCGCCCATTCGCTGAGCT-----CC	23411
Db	5107	TGCTTCGCCGGGGGAGGCGCGCTCGGAACCCGACCGCGCCCATTCGCTGAGCTTCCACCCC	5166
OY	2342	CTTCATATGGGCTCCGTGGTGGCGGCGGACAGCTCCCGGACAGACACACCCCTCTCTCAAGC	2401
Db	5167	CTTCGTGGGCTCCGTGGTGGCGGCGGACAGCTCCCGGACAGACACACCCCTCTCTCAAGC	5226
OY	2402	GCCCAAGTCCCATCAGCACAGCAGAGGGCTGAGAAATGGGGGCGACAGCGACGGGACTGGC	2461
Db	5227	GCCCAAGTCCCATCAGCACACCCAGAGGGCTGAGAAATGGGGGCGACAGCGACGGGACTGGC	5283
OY	2462	AGGCAAGCTCCACCTGCGACGCCCTGGTGGGGAATCCACTGGGTGAAGCCAGCTGGCTCTG	2521
Db	5284	AGGCAAGCTCCACCTGCGACGCCCTGGTGGGGAATCCACTGGGTGAAGCCAGCTGGCTCTG	5342
OY	2522	AGTGTGGTGGAGACTTGGAGAAACCTTTATGTCTAGCTCAGGAGGATCGTAAATACCAATC	2581
Db	5343	AGTGTGGTGGGAGAGTGGAGAACTTTATGTCTAGCTCAGGAGGATTTGAATACCAATC	5402
OY	2582	AGCACCCCTGTGTAGCTCAGGCTGTGTGAATGACCAATTCACACTGTGTATAGCTA	2641
Db	5403	AGCACCCCTGTGTAGCTCAGGCTGTGTGAATGACCAATTCACACTGTGTATAGCTA	5460
OY	2642	CTCGATGGGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGATGTGAATACACCAATCG	2701
Db	5461	CTCGATGGGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGATGTGAATACACCAATCG	5520
OY	2702	GCACCTGTATCTAGCTCAAGGTTTGAACACACCAATCAGCACCTGTGTCTAGCTCA	2761
Db	5521	GCACCTGTATCTAGCTCAAGGTTTGAACACACCAATCAGCACCTGTGTCTAGCTCA	5580
OY	2762	GGGATGTGAGATGACCAATGAGACAGTGTATGTGGCTACTTCATGGG 2811	
Db	5581	GGGATGTGAGATGACCAATGAGACAGTGTATGTGGCTACTTCATGGG 5630	

RESULT 12	
AAZ23901	
ID	AAZ23901 standard; DNA: 49999 BP.
XX	
AC	
XX	AAZ23901;
XX	
DT	25-JAN-2000 (first entry)
XX	
DE	Human LOBO homologue genomic DNA fragment 3.
XX	
KW	LOBO; long bones; bone development; bone extension; skull; osteopathic;
KW	diagnostic; pharmaceutical; gene therapy; transgenic animal; disease;
KW	spondyloepiphyseal dysplasia; achondroplasia; human; ds.
XX	
OS	Homo sapiens.
XX	
BN	W09950284-A2.
XX	
PD	07-OCT-1999.
XX	
PF	26-MAR-1999; 99WO-EP02055.

XX 27-MAR-1998; 98DE-1013799.
PR
XX
XX (ROSE/) ROSENTHAL A.
PA
XX Rosenthal A, Rump A, Hess J, Aligner T, Wirth T;
XX PI
XX DR WPI; 1999-601320/51.
XX
XX Nucleic acids encoding proteins which influence bone development,
PT useful for treating and studying bone disorders -
XX
XX Example 3; Page 272-300; 391pp; German.
XX
XX This invention describes novel nucleic acids (I; designated LOBO (long
CC bones)) encoding proteins influencing bone development in mammals. The
CC proteins of the invention reduce and/or inactivate bone extension (i.e.
CC development), with exception of the skull and have osteopathic activity.
CC The nucleic acid molecules, proteins and antibodies can be used in
CC diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
CC and nucleic acid molecules, etc. are useful for production of transgenic
CC animals, especially a transgenic mouse for the study of diseases
CC associated with bone development, e.g. spondyloepiphyseal dysplasia and
CC achondroplasia. This sequence encodes a human LOBO protein described
CC in the method of the invention.
XX
XX Sequence 49999 BP; 12607 A; 13453 C; 13304 G; 10635 T; 0 other;
SQ

Query Match	11.4%	Score 710.8	DB 20	Length 49999
Best Local Similarity	87.8%	Pred. No.5.9e-149		
Matches	864	Conservative	0	Mismatches 67
				Indels 53
				Gaps 6
QY	1873	TGAGAGCTGACAGATGCCCGGACAGCTCCACAG-CCCTCGTTTCCTCTCGCGGCTCTCTC	1931	
Db	30814	TGAGAGGTGACAGGCTGCTGGCAGTCCCTCACAGGCCCTCGCTTCTCGGACACTCTCTC	3087	
QY	1932	TGCTGGGCTCCACTTGTGTGGCACTTGAGGAGCCCTTCAGCCCAACCCTGCACTGTGG	1991	
Db	30874	TGCTTGGTCTCCCACTTGTGGGCTCACTTGAAGAGACCCCTTCGGCCACCGCTGCACTGTGG	3093	
QY	1992	GAGCCCCCTTGTGGGCTGGCCAAAGGCCAGACCCGGTCTCCCTCACTTTCAGAGGAGGTGTG	2051	
Db	30934	GAGCCCCCTTCTGGGCTGGCCAAAGGCCCGAGACCACTCTCCACTTTCAGAGGAGGTGTG	3099	
QY	2052	GAGGAGAGGCTCAACGAGAAACCGGGGCTGCACAGCGGCTTCGGGGCCAGCTGAGATT	2111	
Db	30994	GAGGAGAGGCGCGCAGCCGGAACCAAGGCTCTCGCGCGGAGCTTCGCGGCCAGCTGAGATT	3105	
QY	2112	CGGGGTGGGCGTGGGGCTTGGCGGGCCCCCGCACTGGAGACAGCGGGCCAGCCCTGCGAGGC	2171	
Db	31054	CGGGGTGGGCTGGGGCTTGGCAGGCCCCCGCACTTGAACAGGCGGGCCGGGCTTCCG-GGC	3111	
QY	2172	CCCGGGCAATGAGAGGCTTTAGCACCCCGGGCCAGCGGCTGCGGAGGGTGTACTGGGGTCCC	2231	
Db	31113	CCCGGGCAATGAGGGGCTTTAGCACCCCGGGCCAGCGGCTGAGAGGGTGTACTGGGGTCCC	3117	
QY	2232	CAGCAGTCCACGCCCGCGGGGCTGTGCTGCTTGATTTCTCACTGGGCTTACCAGCT	2291	
Db	31173	CAGCAGTCCACGACCAACCGGGGTG---CGCTGATTTCTCACCGGGGCTTACTGCTACT	3122	
QY	2292	TCCCGGCGGGCAGGGGCTTCGGGACCTTGCAGCCCGGCATGCTGAGACT-----CCCTTCA	2346	
Db	31229	TCCCGGCGGGGAGGGCTTCGGGACCTTGCAGCCCGGCATGCTGAGACTTCCACCCCTTCA	3128	
QY	2347	TGGGCTCTGTGGGCCCCGAGCTTCGCCGAGACACCAACCCCTGCTCCACAGGCCCCA	2406	
Db	31289	TGGGCTCTGTGGGCCCCGAGCTTCGCCGATGAGACGCGCACCCCTGCTCCATGGCGCCCCA	3134	
QY	2407	GTCCCATGACACGACAAAGGCTTAGAAGTGCAGGGGCGACGAGCAACCGGAACTGGCAGGCA	2466	
Db	31349	GTCCCATGACACCAACCAAGGGCTTAGAGCGTGCAGGGGCGACAG-GGCGGAACTGGCAGGCA	3140	
QY	2467	GCTACCCCTGCGAGCGCTGTGTCGGAAATCCACTGGGTGAAGCCAGACTGGGCTCTCGAGTCT	2526	

Db 31408 GCTCCACCTGCAGCCCGGTCGGAAATCCACTGATAGCGACGTGGCTCCTGAGCTT 31467
QY 2527 GGTGGAGACTTGGAGAACTTTATGCTAGCTCAGGAGGTGAATACAC----- 2576
Db 31468 GGTGGGGCGCTGGAGAACTTTATGCTAGCTCAGGAGTTGTGAATACACCAATCGGCAC 31527
QY 2577 -----CAATCAGCACCTGTGTCTAGCTCAGGCT 2605
Db 31528 TCTGTATCTAGCTCAAGGTTGTAAACACAGACGACACCCCTGTCTAGCTCAGGCT 31587
QY 2606 CTGTGAATGCACCAATCAGACCTGTATCTAGCTAGCTAGTGGGCGCTTGGAGACCT 2665
Db 31588 TTGTGAATGCACCAAGTGCACACCTGTATCTAGCTGCTGTGGGGCCTTGGAGACCT 31647
QY 2666 TTATGCTTACCTCAGGAGTTGTAAATACACCAATCGGACACTGTATCTAGCTCAAGGTT 2725
Db 31648 TTATGCTTACCTCAGGAGTTGTAAATACACCAATCGGACACTGTATCTAGCTCAAGGTT 31707
QY 2726 TGTAAACACACCAATCAGACACCTGTGTCTAGCTCAGGAGTATGTGAATGCACCAATGCAC 2785
Db 31708 TGTAAACACACCAATCAGACACCTGTGTCTAGCTCAGGAGTTGTGAATGCACCAATGCAC 31767
QY 2786 AGTCTGTATCTGGCTACTTTCATG 2809
Db 31768 ACTGTGTATCTAGCTGCTGTGGTG 31791

RESULT 13

AAK67426 standard; DNA; 55235 BP.

AAK67426;

06-NOV-2001 (first entry)

Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:22238.

Human: Immune; haematopoietic; Immune/haematopoietic antigen; cancer;

KW Cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

W0200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 22238; 3071bp + Sequence Listing: English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytoskeletal
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK7694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 55235 BP; 16593 A; 12467 C; 11660 G; 14515 T; 0 other;

Query Match

11.4%; Score 708.4; DB 22; Length 55235;

Best Local Similarity 86.9%; Pred. No. 2,1e-148;
 Matches 857; Conservative 0; Mismatches 76; Indels 53; Gaps 5;

QY	1872	CTGAGAGTGACAGATGCCGGAGTCTCTACAGCCCTGCTGCTGCGGCGCTCTC	1931
DB	45538	CTGAGAGTGACAGATGCCGGAGTCTCTACAGCCCTGCTGCTGCGGCGCTCTC	45597
QY	1932	TGCCTGGGCTCCCACTTCTGGTGGACATTAGAGAGCCCTTACGCCACCCCTGACCTGTG	1991
DB	45598	TGCCTGGGCTCCCACTTCTGGTGGACATTAGAGAGCCCTTACGCCACCCCTGACCTGTG	45657
QY	1992	GAGCCCTTCTGGGCTGGCCAGAGCCAGAGCCGGCTCCCTGAGCTTGGAGAGAGTGG	2051
DB	45658	GAGCCCTTCTGGGCTGGCCAGAGCCAGAGCCGGCTCCCTGAGCTTGGAGAGAGTGG	45717
QY	2052	GAGGAGAGGCTCAGACAGAAACCGGGCTGGCAGCGGCTTGGCGGACCTGAGATT	2111
DB	45718	GAGGAGAGGCTCAGACAGAAACCGGGCTGGCAGCGGCTTGGCGGACCTGAGATT	45777
QY	2112	CCGGTGGGCTGGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2171
DB	45778	CCGGTGGGCTGGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	45835
QY	2172	CCGGGCAATGAGAGGCTTACAGACCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2231
DB	45836	CCGGGCAATGAGAGGCTTACAGACCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	45895
QY	2232	CAGCAGTGCAGCCCGCGGCGGCTGTCTGCTGATTTCTCATGCTGGGCGCTTACAGCCT	2291
DB	45896	CAGCAGTGCAGCCCGCGGCGGCTGTCTGCTGATTTCTCATGCTGGGCGCTTACAGCCT	45951
QY	2292	TCCCGGGGGGAGGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAG	2346
DB	45952	TCCCGGGGGGAGGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAG	46011
QY	2347	TGGGCTCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2406
DB	46012	TGGGCTCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	46071
QY	2407	GTCCATGACACACGAGCGGCTGAGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2466
DB	46072	GTCCATGACACACGAGCGGCTGAGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	46130
QY	2467	GCTACCCCTGAGCCCTGGTGGGAGTCCACTGGGAGGAGCGGCGGCGGCGGCGGCGG	2526
DB	46131	GCTACCCCTGAGCCCTGGTGGGAGTCCACTGGGAGGAGCGGCGGCGGCGGCGGCGG	46190
QY	2527	GGTGGAGACTTGGAGAACCTTATGCTGAGCTCAGGAGATCTAAT-----	2572
DB	46191	GGTGGAGACTTGGAGAACCTTATGCTGAGCTCAGGAGATCTAAT-----	46250
QY	2573	-----ACACCAATCAGACCCCTGTGTACTCAGGGT	2605
DB	46251	-----ACACCAATCAGACCCCTGTGTACTCAGGGT	46310
QY	2606	CTGTGAATGACCAATCCACCTGCTATGCTAGCTGATGAGGGCGCTTGGAGAACT	2665
DB	46311	CTGTGAATGACCAATCCACCTGCTATGCTAGCTGATGAGGGCGCTTGGAGAACT	46370
QY	2666	TATGCTGCTCAGGAGATTTAATATACCAATGCGGCACTGTATCTAGCTCAGGATT	2725
DB	46372	TATGCTGCTCAGGAGATTTAATATACCAATGCGGCACTGTATCTAGCTCAGGATT	46430
QY	2726	TGTAAACACCAATCAGACCCCTGTGTACTCAGGATGATGATGATGATGATGATGATG	2785
DB	46431	TGTAAACACCAATCAGACCCCTGTGTACTCAGGATGATGATGATGATGATGATGATG	46490
QY	2786	AGTGTGATGCTGCTACTTTCATGGG 2811	
DB	46491	AGTGTGATGCTGCTACTTTCATGGG 46516	

RESULT 14

FT	repeat_unit	/tag- j
FT		866..906
FT	repeat_unit	/tag- k
FT		907..943
FT	repeat_unit	/tag- l
FT		944..984
FT	repeat_unit	/tag- m
FT		985..1027
FT	repeat_unit	/tag- n
FT		1028..1066
FT	repeat_unit	/tag- o
FT		1067..1102
FT	repeat_unit	/tag- p
FT		1116..1152
FT	repeat_unit	/tag- q
FT		1153..1193
FT	promoter	/tag- r
FT		1194..1322
FT	misc_feature	/tag- s
FT		1322..1380
FT	misc_feature	/tag- t
FT		/note= "U3 R region"
FT		1381..1680
FT	misc_feature	/tag- u
FT		/note= "U5 region"
FT	repeat_region	/tag- v
FT		/rpt_type= DISPERSED
FT		/note= "3 repeats of 70 bases"
FT	repeat_unit	1380..1452
FT		/tag- w
FT	repeat_unit	1481..1552
FT		/tag- x
FT	repeat_unit	1609..1680
FT		/tag- y
FT	primer_bind	complement (595..616)
FT		/tag- z
FT		/note= "primer of AAA27496 and AA295045"
FT	primer_bind	1677..1694
FT		/tag- aa
FT		/note= "primer of AAA27497"
FT	primer_bind	1700..1726
FT		/tag- ab
FT		/note= "primer of AA295046"
FT	primer_bind	complement (1349..1373)
FT		/tag- ac
FT		/note= "primer of AA295047"
FT	primer_bind	1807..1831
FT		/tag- ad
FT		/note= "primer of AA295048"
FT	LTR	1695..1698
FT		/tag- ae
XX		
PN	MO200023606-A1.	
PD	27-APR-2000.	
XX		
PE	21-OCT-1999;	99MO-US24646.
PR	22-OCT-1998;	98US-0105256.
PA	(MED1-) MEDICAL COLLEGE GEORGIA INST INC.	
PI	Tuan D, Long Q, Bengra C;	
DR	WPI; 2000-339704/29.	
PX		
PT	Retroviral U3 long terminal repeat expression control sequences useful	
PS	for controlling target gene expression, e.g. for gene therapy -	
CC	Claim 1; Fig 2A-B; 76pp; English.	
	The present sequence is that of a solitary endogenous retrovirus 9	

CC (ERV-9) long terminal repeat (LTR) from the boundary area of the
CC human beta-globin locus control region. This ERV-9 LTR contains an
CC unusual U3 enhancer region comprised of 14 tandem repeats with
CC recurrent GATC, CACC and CCAAT motifs (see also AAA27488-92). The
CC LTR is conserved in human and gorilla. In both recombinant
CC constructs and the endogenous human genome, the LTR enhancer and
CC promoter activate the transcription of cis-linked DNA
CC preferentially in erythroid cells. The U3 enhancer, insulator and
CC promoter regions can be used to control expression of any desired
CC gene of interest and used in any vector for this purpose. The
CC enhancer can be used to replace the LTR in retroviral vectors to
CC avoid transcriptional silencing of transgenes and to boost the
CC transcription of a therapeutic transgene in erythroid progenitor
CC cells. The insulators can be used to insulate expression cassettes,
CC especially those to be inserted in the genome of a host cell, from
CC the transcriptional interference and silencing of the flanking
CC host sequences. Thus, the disclosed transcription control
CC elements are preferably used in retroviral vectors to obtain
CC expression in mammalian cells, and especially to express genes in
CC cells in, or to be introduced into, animals (including humans) for
CC gene therapy. They can be also be used in conjunction with any
CC other control sequences and/or vectors.

Query Match	10.7%	Score 669.2;	DB 21;	Length 1831;
Best Local Similarity	68.0%;	Prod. No. 4.1e-140;		
Matches 1396;	Conservative	0;	Mismatches 248;	Indels 408; Gaps 17;
QY	1873	TGAGAGGTGACAGATGCCGCGACAGTCCACAGCCCTGTTGCGCTGCGGCGCCCTCT	1932	
Db	5	TGAGAGGTGACAGAGGTGCTGTGGCAGTCCCTCACAGCCCTGCTCGCTTGTGGGCGCTCTCT	64	
QY	1933	GCTTGGGCTCCCACTTCCGTGTGGCATTGAGAGCCCTTACGCCACCGCTGCATCTGGG	1992	
Db	65	GCTTGGGCTCCCACTTGGTGTGGCATTGAGAGCCCTTACGCCACCGCTGCATCTGTGG	124	
QY	1993	AGCCCTTTCTGGGCTGTGGCCTCAAGGCGACAGCGGCTCCCTACGTTGACGAGAGGTGG	2052	
Db	125	AGCCCTTTCTGGGCTGTGGCCTCAAGGCGACAGCGGCTCCCTACGTTGACGAGAGGTGG	184	
QY	2053	AGGAGAGGCTCAAGACAGAGAACCGGGGCTGCGACAGCGGCTTGGCGGCGACCTGAGTTT	2112	
Db	185	AGGAGACAGACCCGGGCGACAGAAC-CGGCTGTGGCCCGCTGTGAG-----GGAGTTT	235	
QY	2113	CGGGGGGCGTGGGCTTGGCGGGGCGCGCATCGGAGAGGGGCGACGCCCTGCAGGCC	2172	
Db	236	CGGTTGGGCGATGGGCTCCGAGAGACCCGCGCATCGGAGCGCGACGCGGCCACCC	294	
QY	2173	CCGGGCGCATGAGAGGCTTACGACCCCGGCGACGCGGCTGCGGAGGGGTACTGAGGTGCC	2232	
Db	295	GCGGGGAGTGAAGGGCTTACGACCTGGGCGACGAGCTGTCT-----	335	
QY	2233	AGCAGTGCACACCCCGCGGCGCTGTGCTCGCTGCATTTCTACACTGGGCTTACGACCTT	2292	
Db	336	-----TGTCTAAATTCCTTCCGGGCGCTTACGTCCTT	367	
QY	2293	CCGCGGGGCGAGGGCTCGGAGCTGCAGACCCGCGCATGCTGAGGCTCC---CTCTCATGG	2349	
Db	368	CCTCGGGGCGAGGGCTCGGAGCTGCAGCGCGCATGCTGAGGCTCCCGCACCTTCATGG	427	
QY	2350	GCTCTGTGCGGCGCGGAGCTTCCCGGAGAGCACACACCCCTGCTCCACAGCGCCAGTC	2409	
Db	428	GCTCTGTGCGGCGCGGAGCTTCCCGGAGAGCACCGGCGCCCTGCTCCAGGCGACCCAGTC	487	
QY	2410	CCATTCGACACAGCAAGGGCTGAGAAAGTGGCGGGCGACAGCGACCGGAGCTGGCAGGACGT	2469	
Db	488	CCATTCGACACACCAAGGGCTGAABAAGTGGCGGGCC-CAGCAAGGGGACGTGGCAGGACGT	546	
QY	2470	ACCCTTCAGCCCTGTGTGCGAATCCACTGGGTGAAGCCAGTGGGCTCTGAGTCTGGT	2529	
Db	547	CCCCCTCAGGCCAGGTGTGGGAGTCACTGGGTGAACCGCGCTAGG--TCTGAAATTTGCT	605	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2003, 13:22:16 ; Search time 272 Seconds
(without alignments)
10117.735 Million cell updates/sec

Title: US-09-845-020A-5
Perfect score: 6235
Sequence: 1 gatcacttgagagacagtagt.....tgcacttccccacagcct 6235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/1na/PCrus.COMB.seq.*
6: /cgn2_6/ptodata/2/1na/Backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6235	100.0	6235	US-09-305-384-5	Sequence 5, Appl1
2	6235	100.0	6235	US-09-525-160B-6	Sequence 6, Appl1
3	6235	100.0	6679	US-09-305-384-1	Sequence 1, Appl1
4	6235	100.0	6679	US-09-525-160B-5	Sequence 5, Appl1
5	2834	45.5	2834	US-09-305-384-6	Sequence 6, Appl1
6	1252	20.1	1252	US-09-305-384-7	Sequence 7, Appl1
7	669.2	10.7	1831	US-09-422-576D-1	Sequence 1, Appl1
8	563.8	9.0	9704	US-09-814-951A-3	Sequence 3, Appl1
9	462.2	7.4	1091	US-09-422-576D-5	Sequence 5, Appl1
10	442.2	7.1	1165	US-09-422-576D-25	Sequence 25, Appl1
11	439.6	7.1	801	US-09-422-576D-7	Sequence 7, Appl1
12	426.4	6.6	1043	US-09-422-576D-6	Sequence 6, Appl1
13	413.8	6.6	80246	US-09-078-294-4	Sequence 4, Appl1
14	412.2	6.6	80595	US-09-078-294-3	Sequence 3, Appl1
15	401.8	6.4	43950	US-09-735-934A-3	Sequence 3, Appl1
16	401.8	6.4	43950	US-10-060-332-3	Sequence 3, Appl1
17	389.8	6.3	31571	US-08-323-443B-1	Sequence 1, Appl1
18	389.8	6.3	53526	US-08-658-136-2	Sequence 2, Appl1
19	389.8	6.3	53577	US-08-658-136-1	Sequence 1, Appl1
20	381.4	6.1	45716	US-08-965-048-5	Sequence 5, Appl1
21	381.4	6.1	45989	US-08-965-048-6	Sequence 6, Appl1
22	377	6.0	59065	US-09-813-817-3	Sequence 3, Appl1
23	377	6.0	59065	US-09-918-197-3	Sequence 3, Appl1
24	376.2	6.0	51719	US-09-918-686-2	Sequence 2, Appl1
25	376.2	6.0	92139	US-09-918-686-1	Sequence 1, Appl1
26	375.6	6.0	14796	US-08-975-080-35	Sequence 35, Appl1
27	375.6	6.0	14796	US-09-630-706-10	Sequence 10, Appl1

C 28	375.6	6.0	14796	4	US-09-496-694B-3	Sequence 3, Appl1
C 29	366.2	5.9	3885	1	US-08-688-145-1	Sequence 1, Appl1
C 30	366.2	5.9	64467	4	US-09-803-671B-3	Sequence 3, Appl1
C 31	361	5.8	62804	4	US-09-800-960-3	Sequence 3, Appl1
C 32	354.4	5.7	36159	4	US-09-749-588-3	Sequence 3, Appl1
C 33	354.4	5.7	99500	4	US-09-798-098-10	Sequence 10, Appl1
C 34	352.6	5.7	1351	4	US-09-205-258-104	Sequence 104, App
C 35	350.6	5.6	14747	4	US-09-608-285A-42	Sequence 42, Appl
C 36	350.6	5.6	14747	4	US-09-557-800C-42	Sequence 42, Appl
C 37	349.4	5.6	15977	4	US-09-608-285A-59	Sequence 59, Appl
C 38	349.2	5.6	11811	3	US-09-078-294-7	Sequence 7, Appl1
C 39	346.4	5.6	55298	4	US-09-491-356C-1	Sequence 1, Appl1
C 40	341.8	5.5	2713	2	US-08-916-901-6	Sequence 6, Appl1
C 41	341.8	5.5	2713	4	US-09-154-602-6	Sequence 6, Appl1
C 42	340.2	5.5	63588	4	US-09-873-404-3	Sequence 3, Appl1
C 43	338.8	5.4	33060	3	US-08-814-095-7	Sequence 7, Appl1
C 44	338.2	5.4	162450	4	US-09-345-882-1	Sequence 1, Appl1
C 45	336.8	5.4	162450	4	US-09-345-882-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-305-384-5
Sequence 5, Application US/09305384
Patent No. 6242218
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heartlein, Michael W.
APPLICANT: Seldien, Richard F.
TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
FILE REFERENCE: 07236/017001
CURRENT APPLICATION NUMBER: US/09/305,384
EARLIER FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,649
NUMBER OF SEQ ID NOS: 8
SOFTWARE: fastseq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 6235
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-384-5

Query Match 100.0%; Score 6235; DB 3; Length 6235;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GATCCTTGAGAGACAGTACTTCAAGACCGCTGGGACATAGGAGACTCTCTACG	60
DB	1	GATCCTTGAGAGACAGTACTTCAAGACCGCTGGGACATAGGAGACTCTCTACG	60
QY	61	AAAAATCAAAAAATATATGCGCGGATGTCACGTCGTATATCCCTGAACCTTGGG	120
DB	61	AAAAATCAAAAAATATATGCGCGGATGTCACGTCGTATATCCCTGAACCTTGGG	120
QY	121	ACATCAAGCAAGTGTACCTGAGTGTGAGAGTTCAGACTACCTGGCCACATGCT	180
DB	121	ACATCAAGCAAGTGTACCTGAGTGTGAGAGTTCAGACTACCTGGCCACATGCT	180
QY	181	GAACCCATATCTCCACTTAAAAATATCAAAATATACCCAGCATGGTGGCAGGACCTGTA	240
DB	181	GAACCCATATCTCCACTTAAAAATATCAAAATATACCCAGCATGGTGGCAGGACCTGTA	240
QY	241	ATCCCGGCTACTCAGAGAGCTGAGGACAGAGATACCTTGAACCCAGAGGGAGGAGTTG	300
DB	241	ATCCCGGCTACTCAGAGAGCTGAGGACAGAGATACCTTGAACCCAGAGGGAGGAGTTG	300
QY	301	CAGTGAAGTGAATCACCACCTGCATCTCCAGCTGGGTGACAGCAAGACTTATCTC	360
DB	301	CAGTGAAGTGAATCACCACCTGCATCTCCAGCTGGGTGACAGCAAGACTTATCTC	360

QY 361 AAAAAAAAAAAAAAAAAAATTAACGAGCATGTGTGACACCTCTAGTCTCAG 420
 Db 361 AAAAAAAAAAAAAAAAAAATTAACGAGCATGTGTGACACCTCTAGTCTCAG 420
 QY 421 CTACAGAGAGGTGAGGTGAGGATCACTTGAACCTGGGGAGTCAAGGCTACAGTGA 480
 Db 421 CTACAGAGAGGTGAGGTGAGGATCACTTGAACCTGGGGAGTCAAGGCTACAGTGA 480
 QY 481 GCCAAGATCATGCTACACCTCAGAGCTGGGCAACGAGAGAGACCTCTCTAAAAA 540
 Db 481 GCCAAGATCATGCTACACCTCAGAGCTGGGCAACGAGAGAGACCTCTCTAAAAA 540
 QY 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 Db 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 QY 601 ATGTATATAGTTTGGCAACTCAAGATCCAGATAGTCAATTTTATAGGCTTGGGGCGT 660
 Db 601 ATGTATATAGTTTGGCAACTCAAGATCCAGATAGTCAATTTTATAGGCTTGGGGCGT 660
 QY 661 ATGGCTCTGTCCAAATCACTCTGCCCTGTCTTTAGCACAAGGAGCTATAAACAAT 720
 Db 661 ATGGCTCTGTCCAAATCACTCTGCCCTGTCTTTAGCACAAGGAGCTATAAACAAT 720
 QY 721 ACATACATGAAATTTTATAGATGAGATTTGAATTTCAATATGATTTTATATTTAT 780
 Db 721 ACATACATGAAATTTTATAGATGAGATTTGAATTTCAATATGATTTTATATTTAT 780
 QY 781 AAAAAATATCTTTTAAAAAATTTTCCCTTAAACATTTTAAAAAGTAAAGCGGCGCAG 840
 Db 781 AAAAAATATCTTTTAAAAAATTTTCCCTTAAACATTTTAAAAAGTAAAGCGGCGCAG 840
 QY 841 GCCATGTGACGCGCTGTAAATTCAGACACTTTGGAGGCTGAGGTGGGAGATCACTTGA 900
 Db 841 GCCATGTGACGCGCTGTAAATTCAGACACTTTGGAGGCTGAGGTGGGAGATCACTTGA 900
 QY 901 ATCAACAGTTTCGAGACGAGCTGGCCAACTAGCAAAACCCCAATTTCTATAAATAA 960
 Db 901 ATCAACAGTTTCGAGACGAGCTGGCCAACTAGCAAAACCCCAATTTCTATAAATAA 960
 QY 961 AAAAAATAGTGGGCATAGTGTGTCACACCTGTATCCAGCTACTTGGAGGCTGAGGCA 1020
 Db 961 AAAAAATAGTGGGCATAGTGTGTCACACCTGTATCCAGCTACTTGGAGGCTGAGGCA 1020
 QY 1021 GGAGATTCGCTTGAACCTGGGAGCGGAGGTTGCAATGAGCCAACTATGCCACTGAC 1080
 Db 1021 GGAGATTCGCTTGAACCTGGGAGCGGAGGTTGCAATGAGCCAACTATGCCACTGAC 1080
 QY 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTAAAGGCTAT 1140
 Db 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTAAAGGCTAT 1140
 QY 1141 TCCTAATTCAGTGTACATCAGTGTACATCACTGAGTCTGCTGCTGAGGCTAT 1200
 Db 1141 TCCTAATTCAGTGTACATCAGTGTACATCACTGAGTCTGCTGCTGAGGCTAT 1200
 QY 1201 ACCGTGAGGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 1201 ACCGTGAGGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 QY 1261 ACCAAGTTGCATCCAGAGGAGGTTTTTTTTTACAAATCTACACTCCCCCAGCAAAAT 1320
 Db 1261 ACCAAGTTGCATCCAGAGGAGGTTTTTTTTTACAAATCTACACTCCCCCAGCAAAAT 1320
 QY 1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAGAAGGA 1380
 Db 1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAGAAGGA 1380
 QY 1381 AGGGAAGGGAAGTGTGCAAGGCTTTCTTACCATTAAGAAATACCTGGTAGGCTTCT 1440
 Db 1381 AGGGAAGGGAAGTGTGCAAGGCTTTCTTACCATTAAGAAATACCTGGTAGGCTTCT 1440
 QY 1441 GGATGCTGGAAGATGAATAACGGGGGTCTCTGAGAGCTGCCCCCTGTGATGATCAGTGTG 1500

Db 1441 GGATGCTGGAAGATGAATAACGGGGGTCTCTGAGAGCTGCCCTCTGATGATCAGTGTG 1500
 QY 1501 ACTTCTGAGGCTCCAGTCCAGATCCAGGCTGATGATGATGATGATGATGATGATGATG 1560
 Db 1501 ACTTCTGAGGCTCCAGTCCAGATCCAGGCTGATGATGATGATGATGATGATGATGATG 1560
 QY 1561 CACTCTCTGTTTGTGCTTATTTTCCCATGTGGGGCTGAAGTCTGATGATGAGGCTTAT 1620
 Db 1561 CACTCTCTGTTTGTGCTTATTTTCCCATGTGGGGCTGAAGTCTGATGATGAGGCTTAT 1620
 QY 1621 TCAAGATGTACAGCTTTCTTACAGGAAGTGTGTACAGAAACACAGAGGCTTGGCA 1680
 Db 1621 TCAAGATGTACAGCTTTCTTACAGGAAGTGTGTACAGAAACACAGAGGCTTGGCA 1680
 QY 1681 AGATGATCTTAACCTCAATCCATCTGGCTGAGCAGCAGCAGTGTGTGATTTGATTAAC 1740
 Db 1681 AGATGATCTTAACCTCAATCCATCTGGCTGAGCAGCAGCAGTGTGTGATTTGATTAAC 1740
 QY 1741 AAGTTTTTCACTTCTGTGAGGCGATCCCTGGCTACACACACAGTGTGTGACAGGA 1800
 Db 1741 AAGTTTTTCACTTCTGTGAGGCGATCCCTGGCTACACACACAGTGTGTGACAGGA 1800
 QY 1801 TGAATGACGAAGTCCCTTACACTGTAAATCCAGACTTTGGAGGCGCAAGGCGGTGG 1860
 Db 1801 TGAATGACGAAGTCCCTTACACTGTAAATCCAGACTTTGGAGGCGCAAGGCGGTGG 1860
 QY 1861 ATGCGTTGAGCCTGTGAGAGGTGACAGATGCGGCGAGTCTCACAGCCCTGTTGCTCTC 1920
 Db 1861 ATGCGTTGAGCCTGTGAGAGGTGACAGATGCGGCGAGTCTCACAGCCCTGTTGCTCTC 1920
 QY 1921 GGGCGCTCTCTGCTGCTGGCTCCCACTTGGGTGCACTTGAAGAGCCCTTACAGCCACG 1980
 Db 1921 GGGCGCTCTCTGCTGCTGGCTCCCACTTGGGTGCACTTGAAGAGCCCTTACAGCCACG 1980
 QY 1981 CTGCACTGTGGAGGCGCTTCTGCTGGGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTG 2040
 Db 1981 CTGCACTGTGGAGGCGCTTCTGCTGGGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 2041 AGGAGGTGTGAGAGGAGAGGCTTACAGAGAACCGGGGCTGCGACAGGGGCTTGGCGGC 2100
 Db 2041 AGGAGGTGTGAGAGGAGAGGCTTACAGAGAACCGGGGCTGCGACAGGGGCTTGGCGGC 2100
 QY 2101 CAGTGTGAGTTCGGGGTGGGCTGTGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGG 2160
 Db 2101 CAGTGTGAGTTCGGGGTGGGCTGTGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGG 2160
 QY 2161 CCGTGCAGAGCCCGGGCAATGAGAGGCTTACAGCCCGGGGCGGCGGCGGCGGCGGCT 2220
 Db 2161 CCGTGCAGAGCCCGGGCAATGAGAGGCTTACAGCCCGGGGCGGCGGCGGCGGCGGCT 2220
 QY 2221 ACTGGGTGCCAGAGTGCAGGCGCGGCGGCTGTGCTGCTGATTTCTCACTGGGC 2280
 Db 2221 ACTGGGTGCCAGAGTGCAGGCGCGGCGGCTGTGCTGCTGATTTCTCACTGGGC 2280
 QY 2281 CTTAAGCAGCTTCCCGGGGGGAGGAGGCTCGGGAGCTCAGCCCGGCTGCTGAGCTTCC 2340
 Db 2281 CTTAAGCAGCTTCCCGGGGGGAGGAGGCTCGGGAGCTCAGCCCGGCTGCTGAGCTTCC 2340
 QY 2341 CCTCATGAGGCTCCTGTGCGGCGGAGGCTTCCCGAGAGCAGCAGCCCGTGTGCTCAGAG 2400
 Db 2341 CCTCATGAGGCTCCTGTGCGGCGGAGGCTTCCCGAGAGCAGCAGCCCGTGTGCTCAGAG 2400
 QY 2401 CGCCAGTCCCATGACAGCAGAGGAGGCTGAGAAAGTGCAGGCGCAGGCGAGGAGCTGG 2460
 Db 2401 CGCCAGTCCCATGACAGCAGAGGAGGCTGAGAAAGTGCAGGCGCAGGCGAGGAGCTGG 2460
 QY 2461 CAGGAGCTTACCCCTGAGAGGCTGTGTGCGAATTCAGTGGGTGAAGCAGCTGGGCTCT 2520
 Db 2461 CAGGAGCTTACCCCTGAGAGGCTGTGTGCGAATTCAGTGGGTGAAGCAGCTGGGCTCT 2520
 QY 2521 GAGTCTGAGGAGCTTGAAGCCTTTATGTAGTCAAGGAGTGTAAATACCAAT 2580

Dh 2521 GAGTCTGGTGGAGACTTGGAGACCTTTATGTCTAGCTCAGGAGATCGTAATAACCAAT 2580
Qy 2581 CAGACCCCTGTGTAGCTCAGGCTCTGTGTAATGACCAATCCACACTCTGTATCTACT 2640
Db 2581 CAGACCCCTGTGTAGCTCAGGCTCTGTGTAATGACCAATCCACACTCTGTATCTACT 2640
Qy 2641 ACTGTGATGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGATTTGTAATACCAATC 2700
Db 2641 ACTGTGATGGGCGCTTGGAGAACTTTATGTCTAGCTCAGGATTTGTAATACCAATC 2700
Qy 2701 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGCACCCCTGTCTAGCTC 2760
Db 2701 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACACCAATCAGCACCCCTGTCTAGCTC 2760
Qy 2761 AGGCTATGTGATGACCAATCGACAGCTCTATCTGTGCTACTTTCAATGGGCTCCGCT 2820
Db 2761 AGGCTATGTGATGACCAATCGACAGCTCTATCTGTGCTACTTTCAATGGGCTCCGCT 2820
Qy 2821 GAAGAGACCAACAGAGGCTTGTGTGAGCAATTAAGCTCTATCACCCTGGTGCAGGT 2880
Db 2821 GAAGAGACCAACAGAGGCTTGTGTGAGCAATTAAGCTCTATCACCCTGGTGCAGGT 2880
Qy 2881 GGGCTGAGTCCGAAAAGAGAGTCAAGCAAGGAGATTAAGGTGGGCGCTTTTATAGGAT 2940
Db 2881 GGGCTGAGTCCGAAAAGAGAGTCAAGCAAGGAGATTAAGGTGGGCGCTTTTATAGGAT 2940
Qy 2941 TTGGGTAGTAAAGAAAATTTACAGTCAAAAGGGGTTTCTCTGGGCGGAGAGAGT 3000
Db 2941 TTGGGTAGTAAAGAAAATTTACAGTCAAAAGGGGTTTCTCTGGGCGGAGAGAGT 3000
Qy 3001 GGGGTCCCAAGGCTCTAGTGGGGTCTTTTGGCCAGAGATGAGCCAGGAAAAGAGT 3060
Db 3001 GGGGTCCCAAGGCTCTAGTGGGGTCTTTTGGCCAGAGATGAGCCAGGAAAAGAGT 3060
Qy 3061 TTCAAGAGTATATCATCAATTAAGGCAAGACCCGCAATTAACCTCTTTTGAGTG 3120
Db 3061 TTCAAGAGTATATCATCAATTAAGGCAAGACCCGCAATTAACCTCTTTTGAGTG 3120
Qy 3121 GAATGTATCATGTTAAGTTGGGCGAGGCAATTAATCACTCTTTTGTGATTTTCAAGTTAC 3180
Db 3121 GAATGTATCATGTTAAGTTGGGCGAGGCAATTAATCACTCTTTTGTGATTTTCAAGTTAC 3180
Qy 3181 TTCAAGGCAATCTGGGCGTAAATGTGCAATTAAGGAGATGAGGATGGTGGCTGGGCT 3240
Db 3181 TTCAAGGCAATCTGGGCGTAAATGTGCAATTAAGGAGATGAGGATGGTGGCTGGGCT 3240
Qy 3241 CAGAGGCTTGAACACTACTGTGTGGGCGCTTGGAGATGTTTGTGTGACACTCTGTAT 3300
Db 3241 CAGAGGCTTGAACACTACTGTGTGGGCGCTTGGAGATGTTTGTGTGACACTCTGTAT 3300
Qy 3301 CTATGTTAATCTAGTGGGAGCTGGAGAACTTTGTCTAGCTCAGGATTTGTAACGCA 3360
Db 3301 CTATGTTAATCTAGTGGGAGCTGGAGAACTTTGTCTAGCTCAGGATTTGTAACGCA 3360
Qy 3361 CCAATTCAGGCGCTGTCAAAAACAGACACTGCGCTTAACCAATCAGCAGATGAGGAG 3420
Db 3361 CCAATTCAGGCGCTGTCAAAAACAGACACTGCGCTTAACCAATCAGCAGATGAGGAG 3420
Qy 3421 GGCAGATTAAGAGATTAAGAGCAGGCTGCCGAGCCAGCAGCAGTGGCAACGCGCAGAGTCC 3480
Db 3421 GGCAGATTAAGAGATTAAGAGCAGGCTGCCGAGCCAGCAGCAGTGGCAACGCGCAGAGTCC 3480
Qy 3481 CTATCCCAATTAATGGAGCTTTGTCTTTTGTCTGTGGCATTAATCTTGTCTACTGCTCG 3540
Db 3481 CTATCCCAATTAATGGAGCTTTGTCTTTTGTCTGTGGCATTAATCTTGTCTACTGCTCG 3540
Qy 3541 CTTTGTGGTCCACAGCTTTTATAGCTTAACACTCACCAGAAAGGTGTGAGCTTC 3600
Db 3541 CTTTGTGGTCCACAGCTTTTATAGCTTAACACTCACCAGAAAGGTGTGAGCTTC 3600
Qy 3601 ACTCTTAAGGCACTTAAGACAGGCGCACCGGAGAGAAATGAACAATCCGCGCT 3660
Db 3601 ACTCTTAAGGCACTTAAGACAGGCGCACCGGAGAGAAATGAACAATCCGCGCT 3660

Qy 3661 GCGTTAAGAGCTATTAACACTCAGCGGAGAGGCTGACGCTTCACTCTCAGCAGCGAGA 3720
Db 3661 GCGTTAAGAGCTATTAACACTCAGCGGAGAGGCTGACGCTTCACTCTCAGCAGCGAGA 3720
Qy 3721 CCACGAAACCCACAGAAAGAACTCGAACAATCTGAACATCAGAGAGCAAACT 3780
Db 3721 CCACGAAACCCACAGAAAGAACTCGAACAATCTGAACATCAGAGAGCAAACT 3780
Qy 3781 CCAGATGACCACTTAAGAGCTGTAACACTCTGCGAGGGTCCGCGCTTCTTCTTG 3840
Db 3781 CCAGATGACCACTTAAGAGCTGTAACACTCTGCGAGGGTCCGCGCTTCTTCTTG 3840
Qy 3841 AAGTAGAGAGACCAAGACTCCAGTGTGGACACAAAGCCAGAGTTGAGTAGCAGC 3900
Db 3841 AAGTAGAGAGACCAAGACTCCAGTGTGGACACAAAGCCAGAGTTGAGTAGCAGC 3900
Qy 3901 CTGGCAACATGATGAATGCGCTCTGCAAAAAAATAATTAACAAAAATTTAGAGGCGG 3960
Db 3901 CTGGCAACATGATGAATGCGCTCTGCAAAAAAATAATTAACAAAAATTTAGAGGCGG 3960
Qy 3961 AGCATGTGTGCTGCTGCTGTGTGCTCAAGCTACGCGGAGGCTAAAGTGGAGATCGCT 4020
Db 3961 AGCATGTGTGCTGCTGCTGTGTGCTCAAGCTACGCGGAGGCTAAAGTGGAGATCGCT 4020
Qy 4021 TGAGGCTGGAGGCTGAAGACTGCACTGAGCTGTGATTTTACACAGCCCTCTAGGCTGG 4080
Db 4021 TGAGGCTGGAGGCTGAAGACTGCACTGAGCTGTGATTTTACACAGCCCTCTAGGCTGG 4080
Qy 4081 GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAAATTAACAAAAATTAAGAGAGT 4140
Db 4081 GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAAATTAACAAAAATTAAGAGAGT 4140
Qy 4141 GCGTGTATGTGCTAGGCGCAGTGGCTCATGCTGTATCCAGCACTTTGGGAAAGCGAG 4200
Db 4141 GCGTGTATGTGCTAGGCGCAGTGGCTCATGCTGTATCCAGCACTTTGGGAAAGCGAG 4200
Qy 4201 GCGGCGGGGTCACTTAAGCTCAGAGTGTGAGACAGGCTGGCAACATGGAAGAGCC 4260
Db 4201 GCGGCGGGGTCACTTAAGCTCAGAGTGTGAGACAGGCTGGCAACATGGAAGAGCC 4260
Qy 4261 ATCTCTTAAAAAATTAACAAATTTAGCGGCTGTGGGAGAGTGGAGCATGCGCTGTA 4320
Db 4261 ATCTCTTAAAAAATTAACAAATTTAGCGGCTGTGGGAGAGTGGAGCATGCGCTGTA 4320
Qy 4321 TCCAGACTACTCAGAGGCTGAGCAGAGATCACTTGAACCCAGAGGCGCGGTTGC 4380
Db 4321 TCCAGACTACTCAGAGGCTGAGCAGAGATCACTTGAACCCAGAGGCGCGGTTGC 4380
Qy 4381 AGTAGGCGAGATGCTGCATTTGCACTCCACCACCTCCAGGCTGGGCAACAGAGCCAAA 4440
Db 4381 AGTAGGCGAGATGCTGCATTTGCACTCCACCACCTCCAGGCTGGGCAACAGAGCCAAA 4440
Qy 4441 CTCGTCTTAAAAAATAAAGAGGCTGATTAAGAGGTTGCAATGCAATGAG 4500
Db 4441 CTCGTCTTAAAAAATAAAGAGGCTGATTAAGAGGTTGCAATGCAATGAG 4500
Qy 4501 TTGCCAGGCAACATGTTAAGATGTGAGGCTCTGCTTCATATGAGGTTGCAATGAG 4560
Db 4501 TTGCCAGGCAACATGTTAAGATGTGAGGCTCTGCTTCATATGAGGTTGCAATGAG 4560
Qy 4561 CACCTCAAGGCGAGGCTGCAATGCTGATTAATCCAGCACTTTGGAGGCGCAG 4620
Db 4561 CACCTCAAGGCGAGGCTGCAATGCTGATTAATCCAGCACTTTGGAGGCGCAG 4620
Qy 4621 GCGGAGTATCACTGAGAGTCAAGAGTCCAGACAGGCTGCAACCAATGAGGAAT 4680
Db 4621 GCGGAGTATCACTGAGAGTCAAGAGTCCAGACAGGCTGCAACCAATGAGGAAT 4680
Qy 4681 CCAACCTCTACTAAAAAATTAACAAATTTAGATGAGATGATGATGATGATGATGATGAT 4740
Db 4681 CCAACCTCTACTAAAAAATTAACAAATTTAGATGAGATGATGATGATGATGATGATGAT 4740

OY	4741	CTACTTGGGAGCGCTGAGGCGAGAAAATCATTAGAACACGAGGAGCGGAGGTTGTAGTGA	4800
Db	4741	CTACTTGGGAGCGCTGAGGCGAGAAAATCATTAGAACACGAGGAGCGGAGGTTGTAGTGA	4800
OY	4801	CCGAGATCTGTCCATTGTGCATCTCCAGCCTGAGCAATGAGCGAATCCATCTCAAAAAAC	4860
Db	4801	CCGAGATCTGTCCATTGTGCATCTCCAGCCTGAGCAATGAGCGAATCCATCTCAAAAAAC	4860
OY	4861	AACAAACAAAAACCACTCTCTACTCTCCAGGAGCTGGGTACAGACTGTGGCCACATCACT	4920
Db	4861	AACAAACAAAAACCACTCTCTACTCTCCAGGAGCTGGGTACAGACTGTGGCCACATCACT	4920
OY	4921	GCAAGGTCTGAGCGACAGAGCTTAAGGGGGAGCTGACAGAGCCGGGAGCCAGATTAACGTG	4980
Db	4921	GCAAGGTCTGAGCGACAGAGCTTAAGGGGGAGCTGACAGAGCCGGGAGCCAGATTAACGTG	4980
OY	4981	TGTAGATCAGTGTGTGAGATCAGACGTCCTGCCATTGTGTGACACACAGGGGGCCCCA	5040
Db	4981	TGTAGATCAGTGTGTGAGATCAGACGTCCTGCCATTGTGTGACACACAGGGGGCCCCA	5040
OY	5041	AGCACACAGATGTGGCCCCATCCAGTGCACCAATCCATCCATCCAGAGATGTCTGTT	5100
Db	5041	AGCACACAGATGTGGCCCCATCCAGTGCACCAATCCATCCATCCAGAGATGTCTGTT	5100
OY	5101	CTTGGCACGCTGGGGTAAATTAGACAGAAAGTGAACGTCTTGGGTGTGCTCACTACAC	5160
Db	5101	CTTGGCACGCTGGGGTAAATTAGACAGAAAGTGAACGTCTTGGGTGTGCTCACTACAC	5160
OY	5161	TGCCCCAGGCGGCGCTGTGGCGTGTGAAAACGTCAGGGCTGAGGGCGGGCACGGTGGC	5220
Db	5161	TGCCCCAGGCGGCGCTGTGGCGTGTGAAAACGTCAGGGCTGAGGGCGGGCACGGTGGC	5220
OY	5221	TCACGCCCTGTAAATCCACAGCACTTTGGGAGGCGGCGGGGTGATCAAGAGTCAAGAGA	5280
Db	5221	TCACGCCCTGTAAATCCACAGCACTTTGGGAGGCGGCGGGGTGATCAAGAGTCAAGAGA	5280
OY	5281	TCGTGACCATCTGTGGCTTAACACGGGTGAACCCCGCTCTCTACTAAAAATACAAAAATTGG	5340
Db	5281	TCGTGACCATCTGTGGCTTAACACGGGTGAACCCCGCTCTCTACTAAAAATACAAAAATTGG	5340
OY	5341	CCGGGCAATGTGTGGGGGCGACCTGTAGTTCACAGCTACTCGGAGGCTGAGCGAGAGATG	5400
Db	5341	CCGGGCAATGTGTGGGGGCGACCTGTAGTTCACAGCTACTCGGAGGCTGAGCGAGAGATG	5400
OY	5401	GCGTGAACCCGAGAGCGAGATTTTGCAGTAGACCGAGATCGCGCACTGCACCTCAAGCT	5460
Db	5401	GCGTGAACCCGAGAGCGAGATTTTGCAGTAGACCGAGATCGCGCACTGCACCTCAAGCT	5460
OY	5461	GGGGGACAGACAAAGACTCCTATCTGAAAAAATAAAAGAAAGCTTCAGAGTGTGAGCCAGA	5520
Db	5461	GGGGGACAGACAAAGACTCCTATCTGAAAAAATAAAAGAAAGCTTCAGAGTGTGAGCCAGA	5520
OY	5521	GGCCCAAGGCTTAATTCTGTCACTTACCATGACTTGGGCAAGGCACCTCTCCCTGGC	5580
Db	5521	GGCCCAAGGCTTAATTCTGTCACTTACCATGACTTGGGCAAGGCACCTCTCCCTGGC	5580
OY	5581	CCAGTTCACGGGGTTTGAATGAGCTCCAAAGTCCCTTCACACATTAAACGCTGATGGTTC	5640
Db	5581	CCAGTTCACGGGGTTTGAATGAGCTCCAAAGTCCCTTCACACATTAAACGCTGATGGTTC	5640
OY	5641	TAAAGATGAAGATGGGGCACTTCCCTCTCTACCCACAGCCCGGTGTCCACTTCAAGT	5700
Db	5641	TAAAGATGAAGATGGGGCACTTCCCTCTCTACCCACAGCCCGGTGTCCACTTCAAGT	5700
OY	5701	GAATGACCAAGGAAGTCACTGTGTCCCAATCCCGAGTTCCAAAGCCTTGGGGACCTTAC	5760
Db	5701	GAATGACCAAGGAAGTCACTGTGTCCCAATCCCGAGTTCCAAAGCCTTGGGGACCTTAC	5760
OY	5761	TGTAGAGGTCTGCACGAGAGGTGAAGGTGAGGTGAGCCAAATGCGCTCGAAGGGTCTTG	5820
Db	5761	TGTAGAGGTCTGCACGAGAGGTGAAGGTGAGGTGAGCCAAATGCGCTCGAAGGGTCTTG	5820
OY	5821	CCTCATTTGGGACACATCCGGTTTCTCTGGCTTACCGGGATTTGAGGGGCTTTAGC	5880

Db	5821	CCCTAATTCGGGACAGACATCCGGTTTCCTCTGCGCTTACCGGGATTCTAGGGGCTTTAGC	5880
Qy	5881	CGAATGAGATCATGGGGGGCGGGGGGTTTCTGGGGGAGTTCCACGTAATCAACTTGGGA	5940
Db	5881	CGAATGAGATCATGGGGGGCGGGGGGTTTCTGGGGGAGTTCCACGTAATCAACTTGGGA	5940
Qy	5941	CAGGACAGCCCTGGAACCTTTCGATGGTGCCTATCCAAATGTGTGGGTGGGACAGCAACAA	6000
Db	5941	CAGGACAGCCCTGGAACCTTTCGATGGTGCCTATCCAAATGTGTGGGTGGGACAGCAACAA	6000
Qy	6001	GACCCAAATGTCCTTATCTCAGGTAGGGGCTCAGAGGTCTTCCAGACAGGCAGCCTCCGG	6060
Db	6001	GACCCAAATGTCCTTATCTCAGGTAGGGGCTCAGAGGTCTTCCAGACAGGCAGCCTCCGG	6060
Qy	6061	AGAGTTTGGGGGTAGGAATGGGAGCAACACAGGCTTCTTTTCTCTTATGAAATTTGGG	6120
Db	6061	AGAGTTTGGGGGTAGGAATGGGAGCAACACAGGCTTCTTTTCTCTTATGAAATTTGGG	6120
Qy	6121	GGCTTTGGGGGACAGGCTTGGAAATTCCAAAGAGGAGGCAAAAGACATCTCCCAACAAG	6180
Db	6121	GGCTTTGGGGGACAGGCTTGGAAATTCCAAAGAGGAGGCAAAAGACATCTCCCAACAAG	6180
Qy	6181	TCTGCGAGAGGAGAGAGAGAGACCCGACACTGACCTGGCACTTCCCAACAGGCT	6235
Db	6181	TCTGCGAGAGGAGAGAGAGAGACCCGACACTGACCTGGCACTTCCCAACAGGCT	6235

```

RESULT 2
US-09-525-160B-6
; Sequence 6, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evguenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525,160B
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 6235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-6

```

Query Match	100.0%	Score 6235:	DB 4:	Length 6235:
Best Local Similarity	100.0%	Prod. No. 0:		
Matches 6235:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	1	GATCAGCTTGAGAGACAGTAGTTCAAGACACGCTTGCGGAGCATAGGAGAGACTGTCTACAG	60	
DB	1	GATCAGCTTGAGAGACAGTAGTTCAAGACACGCTTGCGGAGCATAGGAGAGACTGTCTACAG	60	
QY	61	AAAAATCAAAAATTAATGCGCGCGGCGATGCTGACGTCCTGTAAATCCCTTAACATTTGGG	120	
DB	61	AAAAATCAAAAATTAATGCGCGGCGCGATGCTGACGTCCTGTAAATCCCTTAACATTTGGG	120	
QY	121	ACATCAAGGCGAAGTGGATCACTTGAAGGTCAGAGAGTCGAGACTGAGCTGAGCCGCGCAACATGGT	180	
DB	121	ACATCAAGGCGAAGTGGATCACTTGAAGGTCAGAGAGTCGAGACTGAGCTGAGCCGCGCAACATGGT	180	
QY	181	GAAACCCCTATCTCCACTTAAAAAATACAAAAATTATGACAGGATGGTGGACAGCACTGTA	240	
DB	181	GAAACCCCTATCTCCACTTAAAAAATACAAAAATTATGACAGGATGGTGGACAGCACTGTA	240	
QY	241	ATCCCGGCTACTCAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGAGGTTG	300	
DB	241	ATCCCGGCTACTCAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGAGGTTG	300	
QY	301	CAGTGAAGTGAATACACACCACTGCACTCCAGCTGGGTGACAGAGCAAGACTTATCTC	360	
DB	301	CAGTGAAGTGAATACACACCACTGCACTCCAGCTGGGTGACAGAGCAAGACTTATCTC	360	

QY	361	AAAAAAAAAAAAAAAAAAAAAAAAAATTACGACGACATGGTAAGTGCACACCTCTAAGTCTCG	420
Db	361	AAAAAAAAAAAAAAAAAAAAAAAAAATTAGCAGGCAATGGTAAGTGCACACCTCTAAGTCTCG	420
QY	421	CTACTCAGAGAGCTGAGGTGGGAGGATCACTTGAACCTGTGGGAGTCAAGGCTACAGTGA	480
Db	421	CTACTCAGAGAGCTGAGGTGGGAGGATCACTTGAACCTGTGGGAGTCAAGGCTACAGTGA	480
QY	481	GCCAAGATCATGCCCCTACACTCCAGCTGGGCCAACAGAGAGACCCGTCTCTAATAA	540
Db	481	GCCAAGATCATGCCCTACACTCCAGCTGGGCCAACAGAGAGACCCGTCTCTAATAA	540
QY	541	AATAATATATATATAAAGAAAAAACAAGCTCTGTTATGTCCTGGTCCATATACATCTAC	600
Db	541	AATAATATATATATAAAGAAAAAACAAGCTCTGTTATGTCCTGGTCCATATACATCTAC	600
QY	601	ATGTATATATAGTTGGCAAACTCAAGATCCAGATAGTCAATTTTTAGGCTGTGGCCGT	660
Db	601	ATGTATATATAGTTGGCAAACTCAAGATCCAGATAGTCAATTTTTAGGCTGTGGCCGT	660
QY	661	ATGGCTCTGTGCACATCACTCTGCCCCTGCTCTTAGACACAAAAGCACTATTAACAT	720
Db	661	ATGGCTCTGTGCACATCACTCTGCCCCTGCTCTTAGACACAAAAGCACTATTAACAT	720
QY	721	ACATACATGAATTTTTTTTATAGACATCGAGATTTGAATTCATATGATTTTTACATTTAT	780
Db	721	ACATACATGAATTTTTTTTATAGACATCGAGATTTGAATTCATATGATTTTTACATTTAT	780
QY	781	AAAAATATCTTTTTAAAAATTTTCCCTTACCATTTTAAAAAGTATAAACCGGCCACGCC	840
Db	781	AAAAATATCTTTTTAAAAATTTTCCCTTACCATTTTAAAAAGTATAAACCGGCCACGCC	840
QY	841	GCCATCGCAGCGCCTGTATATTCACACACTTTTGGGAGGCTGAGGTGGCGAGATCACTTGA	900
Db	841	GCCATCGCAGCGCCTGTATATTCACACACTTTTGGGAGGCTGAGGTGGCGAGATCACTTGA	900
QY	901	ATCAACAGTTGGAGACACAGCCTGGCCAAATAGCAAAACCCCATTTCTACTAAAAATAA	960
Db	901	ATCAACAGTTGGAGACACAGCCTGGCCAAATAGCAAAACCCCATTTCTACTAAAAATAA	960
QY	961	AAAAATATCTGGGCACTAGTGGGACACCTGTGATCCAGCATCTTGGGAGGCTGAGGA	1020
Db	961	AAAAATATCTGGGCACTAGTGGGACACCTGTGATCCAGCATCTTGGGAGGCTGAGGA	1020
QY	1021	GGAGAAATCGCTGGAACCTGGGAGGCGGAGGTGTCAAGTGAACCATCAATGCCACATGCAC	1080
Db	1021	GGAGAAATCGCTGGAACCTGGGAGGCGGAGGTGTCAAGTGAACCATCAATGCCACATGCAC	1080
QY	1081	TTCAGACCTGGGTGACAGAGTGAGACTTGTCTCAACGAAAAAAGTGTAAAGCCAT	1140
Db	1081	TTCAGACCTGGGTGACAGAGTGAGACTTGTCTCAACGAAAAAAGTGTAAAGCCAT	1140
QY	1141	TCCAAATTCAGTGTATCATCAGGTATCAATCACTGAGTTCGGTCAATCTGCTGTGAGGCAT	1200
Db	1141	TCCAAATTCAGTGTATCATCAGGTATCAATCACTGAGTTCGGTCAATCTGCTGTGAGGCAT	1200
QY	1201	ACCTGAGAAGTAGAGTTGCTTGGTCCACAGACATCAATATTTCCACATTAACATAGACAT	1260
Db	1201	ACCTGAGAAGTAGAGTTGCTTGGTCCACAGACATCAATATTTCCACATTAACATAGACAT	1260
QY	1261	ACCAAGTTGCCATCCAGGAGGATTTTTTTTTTATCAATCTACACTCCCCCAGCAACAAT	1320
Db	1261	ACCAAGTTGCCATCCAGGAGGATTTTTTTTTTATCAATCTACACTCCCCCAGCAACAAT	1320
QY	1321	GAGAGTTACTCTCAGATCCTTTACAAAAGATGCTTAAGCCAGTACAGATGAAGACAGA	1380
Db	1321	GAGAGTTACTCTCAGATCCTTTACAAAAGATGCTTAAGCCAGTACAGATGAAGACAGA	1380
QY	1381	AGTGGAGGAGGAGACGTGCGACGCCCTTTAAACCATGAAGAAATACCTGGTAGAGCTTCT	1440
Db	1381	AGTGGAGGAGGAGACGTGCGACGCCCTTTAAACCATGAAGAAATACCTGGTAGAGCTTCT	1440

OY	1441	GGATGCTGGAGATGAATTAACGGGGGCTCTGGAGCTGCCCCCTGTCAATACATCTGTG	1500
Db	1441	GGATGCTGGAGATGAATTAACGGGGGCTCTGGAGCTGCCCCCTGTCAATACATCTGTG	1500
OY	1501	ACTCTGAGCCTCCAGTCCAGTCTCAGCCCCATGTGTCATGGCCAGTATTAATGAGCCCT	1560
Db	1501	ACTCTGAGCCTCCAGTCCAGTCTCAGCCCCATGTGTCATGGCCAGTATTAATGAGCCCT	1560
OY	1561	CACCTCTGTTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTCTGATTGAGCCGTTAT	1620
Db	1561	CACCTCTGTTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTCTGATTGAGCCGTTAT	1620
OY	1621	TCAGATGTACAGCTTTCTTGACAGAAAGTAGTGTCAACAAAACAGAGGGGCTTGGA	1680
Db	1621	TCAGATGTACAGCTTTCTTGACAGAAAGTAGTGTCAACAAAACAGAGGGGCTTGGA	1680
OY	1681	AGATGATTAATCTGCAAAATCTTACCCTGGCTCAGCACAGGTAGTCTGTATCTTGAAC	1740
Db	1681	AGATGATTAATCTGCAAAATCTTACCCTGGCTCAGCACAGGTAGTCTGTATCTTGAAC	1740
OY	1741	AAGTTTTTCACTTCTCTGAGGGCATTCCTTTGGCTACAAACACACAGTTGGTTGACAGA	1800
Db	1741	AAGTTTTTCACTTCTCTGAGGGCATTCCTTTGGCTACAAACACACAGTTGGTTGACAGA	1800
OY	1801	TGAATATGCAAGTCTCTTACACCTGTATCTCCACAGCCTTTGGAGGCCAAGCGGGGTGG	1860
Db	1801	TGAATATGCAAGTCTCTTACACCTGTATCTCCACAGCCTTTGGAGGCCAAGCGGGGTGG	1860
OY	1861	ATGGCTTAGGCTGAGAGTGAACAGACATGCGGGAGTCTCAGACGCCCTGGTTGGCTCTC	1920
Db	1861	ATGGCTTAGGCTGAGAGTGAACAGACATGCGGGAGTCTCAGACGCCCTGGTTGGCTCTC	1920
OY	1921	GGCGGCTCTCTGCGCTTGGGCTCCCACTTGGTGGCAGTTTGAGAGACCTTGAGGCCACCG	1980
Db	1921	GGCGGCTCTCTGCGCTTGGGCTCCCACTTGGTGGCAGTTTGAGAGACCTTGAGGCCACCG	1980
OY	1981	CTGCACCTGTGGAGACCCCTTTCTTGGGCTGGCCAAAGGCCAGCGGCTCCTCAAGCTTGC	2040
Db	1981	CTGCACCTGTGGAGACCCCTTTCTTGGGCTGGCCAAAGGCCAGCGGCTCCTCAAGCTTGC	2040
OY	2041	AGGGAGGTGAGAGGAGAGGCTCAAGCAGGAACCGGGGCGAGCGAGGGGCTTGCGGGGC	2100
Db	2041	AGGGAGGTGAGAGGAGAGGCTCAAGCAGGAACCGGGGCGAGCGAGGGGCTTGCGGGGC	2100
OY	2101	CAGCTGAGAGTTCCGGGTGGGGCTTGGCGGGGCCCGGCACCTCGAGACACGGGGCCAG	2160
Db	2101	CAGCTGAGAGTTCCGGGTGGGGCTTGGCGGGGCCCGGCACCTCGAGACACGGGGCCAG	2160
OY	2161	CCCTGGCAGGCCCGGGGCAATGAGAGGCTTAGACACCGGGCCAGCGGCTGGCGAGGGGT	2220
Db	2161	CCCTGGCAGGCCCGGGGCAATGAGAGGCTTAGACACCGGGCCAGCGGCTGGCGAGGGGT	2220
OY	2221	ACTGGGTGCCCAAGCATGTGCCAGGCCCGCGCGCTGTGCTCGTATTTCTCACTGGGC	2280
Db	2221	ACTGGGTGCCCAAGCATGTGCCAGGCCCGCGCGCTGTGCTCGTATTTCTCACTGGGC	2280
OY	2281	CTTAGCAGGCTTCCCGGGGGGAGGGCTCGGGAGCCTCAGACGCCGCCAATGCTGTAGCCTCC	2340
Db	2281	CTTAGCAGGCTTCCCGGGGGGAGGGCTCGGGAGCCTCAGACGCCGCCAATGCTGTAGCCTCC	2340
OY	2341	CCTCCAGAGGGTCCGTGTGGGGCCGAGCGTCCCGAGAGACCAACCCCGCTGCACAG	2400
Db	2341	CCTCCAGAGGGTCCGTGTGGGGCCGAGCGTCCCGAGAGACCAACCCCGCTGCACAG	2400
OY	2401	CGCCCACTCCCATGACACGACGAAAGGGCTGAAGATGCGGGCCGACGGCACCGGGACTGG	2460
Db	2401	CGCCCACTCCCATGACACGACGAAAGGGCTGAAGATGCGGGCCGACGGCACCGGGACTGG	2460
OY	2461	CAGGCACACTACCCCTGAGAGCCCTGGTCCGGATATCACTGGGTGAAGCAGCTGGGCTCT	2520
Db	2461	CAGGCACACTACCCCTGAGAGCCCTGGTCCGGATATCACTGGGTGAAGCAGCTGGGCTCT	2520
OY	2521	GAGCTCTGTGAGACTTGGAGAAACCTTTATGTCTAGCTCAGGGAGTCTGAATTAACCAAT	2580

|||||
Db 2521 GAGTCTGTGGAGCTTGGAGACCTTTATGTCTAGCTCAGGGATCGTAATATACCAAT 2580
QY 2581 CAGCACCCCTGTCTTACTCAGGGCTGTGTAAATGACACCAATCCACACTGTATCTAGCT 2640
Db 2581 CAGACCCCTGTCTTACTCAGGGCTGTGTAAATGACACCAATCCACACTGTATCTAGCT 2640
QY 2641 ACTCTGTATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATGTAAATATACCAATC 2700
Db 2641 ACTCTGTATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATGTAAATATACCAATC 2700
QY 2701 GGCACCTGTATCTAGCTCAGGCTTGTAAACACCAATCAGACCCCTGTCTTACTC 2760
Db 2701 GGCACCTGTATCTAGCTCAGGCTTGTAAACACCAATCAGACCCCTGTCTTACTC 2760
QY 2761 AGGGTATGTAAATGACACCAATCCAGACTGTCTAGCTACTTTCATTTGGGACATCCGT 2820
Db 2761 AGGGTATGTAAATGACACCAATCCAGACTGTCTAGCTACTTTCATTTGGGACATCCGT 2820
QY 2821 GAAGAGACCAACCAAGGCTTGTGTAGCAATAAAGCTTCTATCCTGGGTGACAGT 2880
Db 2821 GAAGAGACCAACCAAGGCTTGTGTAGCAATAAAGCTTCTATCCTGGGTGACAGT 2880
QY 2881 GGGCTGATCGGAAAGAGTACGCAAGGAGATAAAGGTGGGGCTTTTATAGAT 2940
Db 2881 GGGCTGATCGGAAAGAGTACGCAAGGAGATAAAGGTGGGGCTTTTATAGAT 2940
QY 2941 TTGGGTAGGTAAAGGAAATTTACAGTCAAGAGGGGTTGTCTCTGGCGGACAGAGTGG 3000
Db 2941 TTGGGTAGGTAAAGGAAATTTACAGTCAAGAGGGGTTGTCTCTGGCGGACAGAGTGG 3000
QY 3001 GGGGTGCGCAAGGTCTAGTGGGGGCTTTTGTAGCCAGAGATAGCCAGGAAAGAGACT 3060
Db 3001 GGGGTGCGCAAGGTCTAGTGGGGGCTTTTGTAGCCAGAGATAGCCAGGAAAGAGACT 3060
QY 3061 TTCAAGAGTATGTCTCATATTAAAGCAAGGACCGGCATTTACACTCTTTTGTGGT 3120
Db 3061 TTCAAGAGTATGTCTCATATTAAAGCAAGGACCGGCATTTACACTCTTTTGTGGT 3120
QY 3121 GAATGTCTATGATTAAGTTGGGGCAGGGCATATTCACTTCTTTTGTATCTTCAATTAC 3180
Db 3121 GAATGTCTATGATTAAGTTGGGGCAGGGCATATTCACTTCTTTTGTATCTTCAATTAC 3180
QY 3181 TTGAGGCACTCTGGGCTATATGTGCAAGTTACAGGGGATGCCATGGCTTGGGCT 3240
Db 3181 TTGAGGCACTCTGGGCTATATGTGCAAGTTACAGGGGATGCCATGGCTTGGGCT 3240
QY 3241 CAGAGGCTTACAGCTACTCTGGTGGGGCTTGGAGAAATGTTTGTGCACTCTGTAT 3300
Db 3241 CAGAGGCTTACAGCTACTCTGGTGGGGCTTGGAGAAATGTTTGTGCACTCTGTAT 3300
QY 3301 CTATGTTATCTAGTGGGACGTGGAGAACCTTGTGTCTAGCTCAGGGATGTAAAGCA 3360
Db 3301 CTATGTTATCTAGTGGGACGTGGAGAACCTTGTGTCTAGCTCAGGGATGTAAAGCA 3360
QY 3361 CCAATCGGCGCTGTAAACAGACACCTCGGCTCTACCAATCAGAGATGTGGGTCG 3420
Db 3361 CCAATCGGCGCTGTAAACAGACACCTCGGCTCTACCAATCAGAGATGTGGGTCG 3420
QY 3421 GGGCAGATAGAGAAATTAAGAGAGGTGCCGAGCCAGAGTGGCAACGGCACAGTTC 3480
Db 3421 GGGCAGATAGAGAAATTAAGAGAGGTGCCGAGCCAGAGTGGCAACGGCACAGTTC 3480
QY 3481 CTATCCAAATATGGCAGCTTGTCTTTGTCTTTGCGATAAATCTTCTACTGCTCG 3540
Db 3481 CTATCCAAATATGGCAGCTTGTCTTTGTCTTTGCGATAAATCTTCTACTGCTCG 3540
QY 3541 CTTTGGGTGACACAGCTTTTATGAGCTGTAAACATCAGCAAGAGGCTCTGAGCTTC 3600
Db 3541 CTTTGGGTGACACAGCTTTTATGAGCTGTAAACATCAGCAAGAGGCTCTGAGCTTC 3600
QY 3601 ACTCTGTAAGCACTAAGACACGAGCCACGGGAGAGATGAACAACCTCGGGCGGCT 3660
|||||

Db 3601 ACTCTGTAAGCACTAAGACACGACGCCACCGGAGAGATGAACAACCTCGGGCGGCT 3660
QY 3661 GCTTTAAGAGCTATTAACACTACCGCAAGGCTGTGAGCTTCACTCTCCACGACGAGA 3720
Db 3661 GCTTTAAGAGCTATTAACACTACCGCAAGGCTGTGAGCTTCACTCTCCACGACGAGA 3720
QY 3721 CCACGAACCCACGAAGGAAGAACTGCGAACATCTGAAATCAGAAAGAACT 3780
Db 3721 CCACGAACCCACGAAGGAAGAACTGCGAACATCTGAAATCAGAAAGAACT 3780
QY 3781 CCAGATCACCACCTTAAGAGCTGTAAACATCACTGTGAGGGTCCGGGCTCTCTTG 3840
Db 3781 CCAGATCACCACCTTAAGAGCTGTAAACATCACTGTGAGGGTCCGGGCTCTCTTG 3840
QY 3841 AAGTCACTGAGACCAACCACTACACAGTTTCGACACAGCCAGAGATTGAGATCAGC 3900
Db 3841 AAGTCACTGAGACCAACCACTACACAGTTTCGACACAGCCAGAGATTGAGATCAGC 3900
QY 3901 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGG 3960
Db 3901 CTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGG 3960
QY 3961 AGCATGGTGGTCCGCTGTGGTCCAGCTACCGGGGAGGCTAAAGTGGAGATCGCT 4020
Db 3961 AGCATGGTGGTCCGCTGTGGTCCAGCTACCGGGGAGGCTAAAGTGGAGATCGCT 4020
QY 4021 TGAGCCCTGGGAGTGAAGACTGCAAGTGAAGTGTATGACACAGCCCTCTAGCTGGG 4080
Db 4021 TGAGCCCTGGGAGTGAAGACTGCAAGTGAAGTGTATGACACAGCCCTCTAGCTGGG 4080
QY 4081 GGCAGACTGAGACCTGTTTCCCTCCGCAAAAAATTTGACAAAGTGTAAAGGT 4140
Db 4081 GGCAGACTGAGACCTGTTTCCCTCCGCAAAAAATTTGACAAAGTGTAAAGGT 4140
QY 4141 GCTGTATGTGGTGGGCGCATGGCTCATGGCTTATCCAGACTTTGGGAAGCGAG 4200
Db 4141 GCTGTATGTGGTGGGCGCATGGCTCATGGCTTATCCAGACTTTGGGAAGCGAG 4200
QY 4201 GCGGGCGGGTCACTTAAGGTGAGAGTGTGAGACAGCCTGGCCACATGAGAAAGCCC 4260
Db 4201 GCGGGCGGGTCACTTAAGGTGAGAGTGTGAGACAGCCTGGCCACATGAGAAAGCCC 4260
QY 4261 ATCTCTTCTAAAAATTAACAAATTTAGCGGCTGTGGGGCAGTGGTGAAGCATGCTTAA 4320
Db 4261 ATCTCTTCTAAAAATTAACAAATTTAGCGGCTGTGGGGCAGTGGTGAAGCATGCTTAA 4320
QY 4321 TCCAGCTACTCAGAGGCTGAGGACGAGAGAACTTGAACCCAGAGGGGGGCTTGC 4380
Db 4321 TCCAGCTACTCAGAGGCTGAGGACGAGAGAACTTGAACCCAGAGGGGGGCTTGC 4380
QY 4381 AGTGAGCCGAGATCGTCATTTGCACTCCACCCACTCAGCTGGGCAACAGAGCCAAA 4440
Db 4381 AGTGAGCCGAGATCGTCATTTGCACTCCACCCACTCAGCTGGGCAACAGAGCCAAA 4440
QY 4441 CTCTGTCTTAAAAAATTTAAAAAATTTGCTGTGACATTAAGAGGTGTGCAATGATAG 4500
Db 4441 CTCTGTCTTAAAAAATTTAAAAAATTTGCTGTGACATTAAGAGGTGTGCAATGATAG 4500
QY 4501 TTGCGAGGCAACATGTTAAGAAATGTGAGCTCTGCTTCATATGAGCTGTAAAAAC 4560
Db 4501 TTGCGAGGCAACATGTTAAGAAATGTGAGCTCTGCTTCATATGAGCTGTAAAAAC 4560
QY 4561 CACCCCTAAGGCCAGGTGAGTGGCTCATGTATATCCACAGACTTTGGAGGCGGAG 4620
Db 4561 CACCCCTAAGGCCAGGTGAGTGGCTCATGTATATCCACAGACTTTGGAGGCGGAG 4620
QY 4621 GCGGGTGTGATCAGCTGAGGTGAGAGTTCAGAGCACCTGACCAACCATGTGTAAAT 4680
Db 4621 GCGGGTGTGATCAGCTGAGGTGAGAGTTCAGAGCACCTGACCAACCATGTGTAAAT 4680
QY 4681 CCCAGCTCTACTAAAAATTAACAAATTTAGATGAGAGATGGTGTGATGCTTATCCAC 4740
Db 4681 CCCAGCTCTACTAAAAATTAACAAATTTAGATGAGAGATGGTGTGATGCTTATCCAC 4740

4741 CTACTTGGAGGCTGAGGAGGAAATCATCTAGAACAGGAGGAGGTTGTGTGAG 4800
4741 CTACTTGGAGGCTGAGGAGGAAATCATCTAGAACAGGAGGAGGTTGTGTGAG 4800
4801 CCGAGATGCTGACATTCACCTCAGCCTGAGCAATGAGCAAACTCTCATCTCAAAAAAC 4860
4801 CCGAGATGCTGACATTCACCTCAGCCTGAGCAATGAGCAAACTCTCATCTCAAAAAAC 4860
4861 AACAAACAAACCCCTCTCTACTCCAGGAGAGCTGGGTACAGAGCTGGCCACATCACT 4920
4861 AACAAACAAACCCCTCTCTACTCCAGGAGAGCTGGGTACAGAGCTGGCCACATCACT 4920
4921 GCAAGGTCTGAGGACACAGAGCTAAGGCGGAGCTCAGAGACCGGAGACAGTAACAGG 4980
4921 GCAAGGTCTGAGGACACAGAGCTAAGGCGGAGCTCAGAGACCGGAGACAGTAACAGG 4980
4981 TGTGAGATCACTGTGTGAGATCAGACGTCCTGCCATTGGTACACCAAGGGGCCCCCA 5040
4981 TGTGAGATCACTGTGTGAGATCAGACGTCCTGCCATTGGTACACCAAGGGGCCCCCA 5040
5041 AGCACCAGAGATGGCCCATCAGTACACATCCATCTTCATCCAGAGATGTCTGTTT 5100
5041 AGCACCAGAGATGGCCCATCAGTACACATCCATCTTCATCCAGAGATGTCTGTTT 5100
5101 CTGGCAGCTGGGGTAATTAAGACAGAGAGTGAAGTCTGGGTGTGTGAGTCAAGAC 5160
5101 CTGGCAGCTGGGGTAATTAAGACAGAGAGTGAAGTCTGGGTGTGTGAGTCAAGAC 5160
5161 TGCCCCAGGAGGCTTGTGGCTGTAGAAAAAGTTACAGGCTTACGCGGAGCGGTGGC 5220
5161 TGCCCCAGGAGGCTTGTGGCTGTAGAAAAAGTTACAGGCTTACGCGGAGCGGTGGC 5220
5221 TCAGGCGGTATATCCACACTTTGGGAGCGGAGGCGGTGATCAGAGGTGAGAGA 5280
5221 TCAGGCGGTATATCCACACTTTGGGAGCGGAGGCGGTGATCAGAGGTGAGAGA 5280
5281 TCAGGCGGTATATCCACACTTTGGGAGCGGAGGCGGTGATCAGAGGTGAGAGA 5280
5281 TCAGGCGGTATATCCACACTTTGGGAGCGGAGGCGGTGATCAGAGGTGAGAGA 5280
5341 CCGGCGAGCTGGGCTTAACAGGAGTAAACCCCTCTCTACTAAAAATACAAAAATTTGG 5340
5341 CCGGCGAGCTGGGCTTAACAGGAGTAAACCCCTCTCTACTAAAAATACAAAAATTTGG 5340
5341 CCGGCGAGCTGGGCTTAACAGGAGTAAACCCCTCTCTACTAAAAATACAAAAATTTGG 5340
5341 CCGGCGAGCTGGGCTTAACAGGAGTAAACCCCTCTCTACTAAAAATACAAAAATTTGG 5340
5341 CCGGCGAGCTGGGCTTAACAGGAGTAAACCCCTCTCTACTAAAAATACAAAAATTTGG 5340
5401 GCGTGAACCCGAGAGGAGAGTGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 5460
5401 GCGTGAACCCGAGAGGAGAGTGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 5460
5401 GCGTGAACCCGAGAGGAGAGTGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 5460
5461 GGGGCGAGAGCAAGACTTCATCTGGAAGAAAAAGAAAAAGTTCAAGTGTGAGCCAGA 5520
5461 GGGGCGAGAGCAAGACTTCATCTGGAAGAAAAAGAAAAAGTTCAAGTGTGAGCCAGA 5520
5461 GGGGCGAGAGCAAGACTTCATCTGGAAGAAAAAGAAAAAGTTCAAGTGTGAGCCAGA 5520
5521 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5521 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5581 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5581 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5581 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5581 GGGGCGAGCTGTAATTTCTGCTACCTTACCATGACCTTGGGCAAGGCACTTCCTTCC 5580
5641 TAAATGAGAGATGGGAGTGGGCTTCCCTCTCTACCCAGCCGCTGTCACCTTCAAGT 5700
5641 TAAATGAGAGATGGGAGTGGGCTTCCCTCTCTACCCAGCCGCTGTCACCTTCAAGT 5700
5641 TAAATGAGAGATGGGAGTGGGCTTCCCTCTCTACCCAGCCGCTGTCACCTTCAAGT 5700
5641 TAAATGAGAGATGGGAGTGGGCTTCCCTCTCTACCCAGCCGCTGTCACCTTCAAGT 5700
5701 GAATGAGAGAGTCACTGTCTCCATCCCGAGTTCACAAAGCCCTTGGGAGACCTTAC 5760
5701 GAATGAGAGAGTCACTGTCTCCATCCCGAGTTCACAAAGCCCTTGGGAGACCTTAC 5760
5701 GAATGAGAGAGTCACTGTCTCCATCCCGAGTTCACAAAGCCCTTGGGAGACCTTAC 5760
5701 GAATGAGAGAGTCACTGTCTCCATCCCGAGTTCACAAAGCCCTTGGGAGACCTTAC 5760
5761 TGTGAGGCTGTGACAGAGAGTGAAGTCAAGTGAAGTCAAGTGAAGTGAAGTGAAGT 5820
5761 TGTGAGGCTGTGACAGAGAGTGAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5820
5761 TGTGAGGCTGTGACAGAGAGTGAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5820

5821 CCTCATTCGGGACAGACATCCGTTTCTCTGCTCTACCCGGGATTTAGGGGCTTTAGC 5880
5821 CCTCATTCGGGACAGACATCCGTTTCTCTGCTCTACCCGGGATTTAGGGGCTTTAGC 5880
5881 CGAATGAGTCAATGGGGGCGGGGGGTTCTGGGGAGTTCACCATTAATCAATTTGGGA 5940
5881 CGAATGAGTCAATGGGGGCGGGGGGTTCTGGGGAGTTCACCATTAATCAATTTGGGA 5940
5941 CAGGACAGCTTGAATTTCTCATGTGCTCTATCCAAAGTGTGGGTGGGACAGCCAA 6000
5941 CAGGACAGCTTGAATTTCTCATGTGCTCTATCCAAAGTGTGGGTGGGACAGCCAA 6000
6001 GACCAATGCTCTTATCTCAGTGAAGGCTCAGAGGCTCCAGACAGGAGCTCCGG 6060
6001 GACCAATGCTCTTATCTCAGTGAAGGCTCAGAGGCTCCAGACAGGAGCTCCGG 6060
6061 AGAGTTGGGGTAGAATGGAGAACCCAGGCTTCTTTTCTCTCTTGAATTTGGG 6120
6061 AGAGTTGGGGTAGAATGGAGAACCCAGGCTTCTTTTCTCTCTTGAATTTGGG 6120
6121 GGGTGGGGGAGAGCTTGAAGTCCAAAGAGAGGGGCAAGACATCCCGACAG 6180
6121 GGGTGGGGGAGAGCTTGAAGTCCAAAGAGAGGGGCAAGACATCCCGACAG 6180
6181 TCTGCCAGAGGAGAGAGGAGAGCCCGACTCAGTGCCTTCCCAAGGCT 6235
6181 TCTGCCAGAGGAGAGAGGAGAGCCCGACTCAGTGCCTTCCCAAGGCT 6235
6181 TCTGCCAGAGGAGAGAGGAGAGCCCGACTCAGTGCCTTCCCAAGGCT 6235

RESULT 3
US-09-305-384-1
; Sequence 1, Application US/09305384
; Patent No. 6242218
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
; FILE REFERENCE: 07236/017001
; CURRENT APPLICATION NUMBER: US/09/305,384
; EARLIER FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,649
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 6679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-305-384-1
Query Match 100.0%; Score 6235; DB 3; Length 6679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GATCACTGAGGACAGTAGTTCAGAACCAAGCCTGGGACGATAGGAGACGTCTCTACG 60
20 GATCACTGAGGACAGTAGTTCAGAACCAAGCCTGGGACGATAGGAGACGTCTCTACG 79
61 AAAAATCAAAAAATTTATGGCCGGCATGGTGGCTCAGCTGTATATCCCTGAACCTTTGGG 120
80 AAAAATCAAAAAATTTATGGCCGGCATGGTGGCTCAGCTGTATATCCCTGAACCTTTGGG 139
121 ACATCAAGGCAAGTGGATCACTTGAAGTTCAGAGTTCGAGACTACCTGGCCACATGCT 180
140 ACATCAAGGCAAGTGGATCACTTGAAGTTCAGAGTTCGAGACTACCTGGCCACATGCT 199
181 GAAACCCATCTCTCACTAAAAATACAAAAATTTACCCAGGAGTGTGGTGGACGACCTGTA 240
200 GAAACCCATCTCTCACTAAAAATACAAAAATTTACCCAGGAGTGTGGTGGACGACCTGTA 259
241 ATCCGGGTCTACAGAGAGCTGAGGACAGAGATCACTGTAACCCAGAGGCGGAGTTTG 300

Dh 260 ATCCCGGCTACTGAGGCGCTGAGGCGAGAGATCACTTGAAACCAGAGGCGGAGGTTG 319
Qy 301 CAGTAGCTGAGATCACACCACCTGCTCCAGCCTGGGTGACAGAGCAAGACTTATCTC 360
Dh 320 CAGTAGCTGAGATCACACCACCTGCTCCAGCCTGGGTGACAGAGCAAGACTTATCTC 379
Qy 361 AAAAAAAAAAAAAAAAAAATTTAGCCAGGCTATGTGTGACACCTCTAGCTTCAG 420
Dh 380 AAAAAAAAAAAAAAAAAAATTTAGCCAGGCTATGTGTGACACCTCTAGCTTCAG 439
Qy 421 CTACTCGAGGCGTGAAGTGGAGGATCACTTGAACCTGGGGCGGTGAAGGCTCACTGA 480
Dh 440 CTACTCGAGGCGTGAAGTGGAGGATCACTTGAACCTGGGGCGGTGAAGGCTCACTGA 499
Qy 481 GCCAAGATCATGCCACTACATCCAGCCTGGGCAACAGAGAGAGACCCTCTCTATAAAA 540
Dh 500 GCCAAGATCATGCCACTACATCCAGCCTGGGCAACAGAGAGAGACCCTCTCTATAAAA 559
Qy 541 AATTAATTAATTAAGAAAAAACAGCTGTGTTATGTCTCTGCTGCTCAATCACTACT 600
Dh 560 AATTAATTAATTAAGAAAAAACAGCTGTGTTATGTCTCTGCTGCTCAATCACTACT 619
Qy 601 ATGATATGTTGGCAAACTCAAGATCCAGATGATCAATTTTGTAGGCTTGGGGCGT 660
Dh 620 ATGATATGTTGGCAAACTCAAGATCCAGATGATCAATTTTGTAGGCTTGGGGCGT 679
Qy 661 ATGCTCTGTCACAATCACTGCTGCTCTCTTCTAGCACAAAGCAGCTATTAACAT 720
Dh 680 ATGCTCTGTCACAATCACTGCTGCTCTCTTCTAGCACAAAGCAGCTATTAACAT 739
Qy 721 ACATACATGAATTTTATATAGACATGAGATTTGAATTCATATGATTTTACATTTTAT 780
Dh 740 ACATACATGAATTTTATATAGACATGAGATTTGAATTCATATGATTTTACATTTTAT 799
Qy 781 AAAAAAAAAAAAAAAAAATTTCCCTAACATTTAAAGTGAAGGCGGCGCACGCG 840
Dh 800 AAAAAAAAAATTTTAAAAATTTTCCCTAACATTTAAAGTGAAGGCGGCGCACGCG 859
Qy 841 GCCATGCTCACGCGCTGTAATTCAGCAGCTTGGAGGCTGAGTGGGCGAGATCACTTGAG 900
Dh 860 GCCATGCTCACGCGCTGTAATTCAGCAGCTTGGAGGCTGAGTGGGCGAGATCACTTGAG 919
Qy 901 ATCAACGTTGAGAGACCAGCCTGGCCAACTAGCAAAACCCCATTTTCTACTAAAAATAAA 960
Dh 920 ATCAACGTTGAGAGACCAGCCTGGCCAACTAGCAAAACCCCATTTTCTACTAAAAATAAA 979
Qy 961 AAAAAAAAAAGCTGAGTGTGTCACACCTGTGATCCAGCTACTTGGAGGCTGAGGCA 1020
Dh 980 AAAAAAAAAAGCTGAGTGTGTCACACCTGTGATCCAGCTACTTGGAGGCTGAGGCA 1039
Qy 1021 GGAAGATCGCTGAAACCTGGGAAACGGAGGTTGAGAGTGAAGCAACATCATGCGACTGCAC 1080
Dh 1040 GGAAGATCGCTGAAACCTGGGAAACGGAGGTTGAGAGTGAAGCAACATCATGCGACTGCAC 1099
Qy 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTGTAAAAAGCAT 1140
Dh 1100 TCCAGCCTGGGTGACAGAGTGAAGTCTGCTCAACGAAAAAAGTGTAAAAAGCAT 1159
Qy 1141 TCCTAATTCAGTATCATCTAGTATCATCTAGGTTGCTGCTACTCTGCTCTGAGGCA 1200
Dh 1160 TCCTAATTCAGTATCATCTAGTATCATCTAGGTTGCTGCTACTCTGCTCTGAGGCA 1219
Qy 1201 ACCTGAGAGTGAAGTGTGCTGCTGACAGAGATCAATCAATTTCCATTAATCAAGACAT 1260
Dh 1220 ACCTGAGAGTGAAGTGTGCTGCTGACAGAGATCAATCAATTTCCATTAATCAAGACAT 1279
Qy 1261 ACCAAGTGCATCCAAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Dh 1280 ACCAAGTGCATCCAAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1339
Qy 1321 GAGAGTGTCTCAGATCTCTTACAAAGATGCTCTAAGCCAGTACAGATGATAAAGAGA 1380
Dh 1340 GAGAGTGTCTCAGATCTCTTACAAAGATGCTCTAAGCCAGTACAGATGATAAAGAGA 1399

Qy 1381 AGTGGAGGGGAAAGCTCCAGCCCTTCTAACCATGAAGAAATACCTGTAGAGCCCTTCT 1440
Dh 1400 AGTGGAGGGGAAAGCTCCAGCCCTTCTAACCATGAAGAAATACCTGTAGAGCCCTTCT 1459
Qy 1441 GGATGCTGGAAGGATGAATTAACGGGGGCTCTGTGAGGCTCCGCCCTCTCAGATCACTGTG 1500
Dh 1460 GGATGCTGGAAGGATGAATTAACGGGGGCTCTGTGAGGCTCCGCCCTCTCAGATCACTGTG 1519
Qy 1501 ACTTCTGAGCCTCCAGTCCAGTCTCAGCCCCCATGTGTCATGGGCGAGTGAATGAGCCCT 1560
Dh 1520 ACTTCTGAGCCTCCAGTCCAGTCTCAGCCCCCATGTGTCATGGGCGAGTGAATGAGCCCT 1579
Qy 1561 CACTCTGTGTTGGTCTTATTTCTCCCATGTGGGGCTGAAGTCTGATTAAGCCGTTAT 1620
Dh 1580 CACTCTGTGTTGGTCTTATTTCTCCCATGTGGGGCTGAAGTCTGATTAAGCCGTTAT 1639
Qy 1621 TCAAGATGTACAGCTTTCTTGACAGGAAGTAGTGTACAGAAACAGAGGGGCTTGGA 1680
Dh 1640 TCAAGATGTACAGCTTTCTTGACAGGAAGTAGTGTACAGAAACAGAGGGGCTTGGA 1699
Qy 1681 AGATGATCTAATCTGCAAAATCCTACCTGGCTCAGGCAACAGTGTGATCTTGATAC 1740
Dh 1700 AGATGATCTAATCTGCAAAATCCTACCTGGCTCAGGCAACAGTGTGATCTTGATAC 1759
Qy 1741 AAGTTTTTCACTTCTCTGAGGCGCATCCCTGGCTACACACACACATTTGGTTGACAGGA 1800
Dh 1760 AAGTTTTTCACTTCTCTGAGGCGCATCCCTGGCTACACACACACATTTGGTTGACAGGA 1819
Qy 1801 TGAAGATGACGAATCCCTTACACCTGTATATCCAGCAGCTTTGGAGGCGCAAGGGGCTGG 1860
Dh 1820 TGAAGATGACGAATCCCTTACACCTGTATATCCAGCAGCTTTGGAGGCGCAAGGGGCTGG 1879
Qy 1861 ATGGCTGTAGCTGAGAGGTGACAGCATGCGGCGAGTCCACAGCCCTGTGGCTGCTC 1920
Dh 1880 ATGGCTGTAGCTGAGAGGTGACAGCATGCGGCGAGTCCACAGCCCTGTGGCTGCTC 1939
Qy 1921 GGGCGCTCTCTGCTGGGCTGCCACTTGGGTGACATTTGAGAGAGCCCTTCAAGCCACCG 1980
Dh 1940 GGGCGCTCTCTGCTGGGCTGCCACTTGGGTGACATTTGAGAGAGCCCTTCAAGCCACCG 1999
Qy 1981 CTGCACGTGTGGAGCCCTTTCTGAGGCTGGGCGAAGGCCAAGCGGCTCCCTCAGCTTGC 2040
Dh 2000 CTGCACGTGTGGAGCCCTTTCTGAGGCTGGGCGAAGGCCAAGCGGCTCCCTCAGCTTGC 2059
Qy 2041 AGGAGGTGTGAGAGGAGGAGGCTCAACAGCAAGAACGGGGCTGGCGACAGGCGCTTGCGGGC 2100
Dh 2060 AGGAGGTGTGAGAGGAGGAGGCTCAACAGCAAGAACGGGGCTGGCGACAGGCGCTTGCGGGC 2119
Qy 2101 CAGCTGAGTTCGGGCTGGGCTGGGCTTGGGCGGCCCGCACTCGGAGCAGGCGCAG 2160
Dh 2120 CAGCTGAGTTCGGGCTGGGCTGGGCTTGGGCGGCCCGCACTCGGAGCAGGCGCAG 2179
Qy 2161 CCCTGCCAGGCCCGCGGCAATGAGAGGCTTAGACACCGGGGCGAGGGGCTCGGAGGGTGT 2220
Dh 2180 CCCTGCCAGGCCCGCGGCAATGAGAGGCTTAGACACCGGGGCGAGGGGCTCGGAGGGTGT 2239
Qy 2221 ACTGGGTGCCCAAGCAGTGTGCAAGCCCGGGCTGTGCTCGCTGCAATTTCTACTAGGGC 2280
Dh 2240 ACTGGGTGCCCAAGCAGTGTGCAAGCCCGGGCTGTGCTCGCTGCAATTTCTACTAGGGC 2299
Qy 2281 CTTAGCAGCTTCCCGGGGCGAGGGCTGCGGACCTTGACGCCCGCATGAGCTTCC 2340
Dh 2300 CTTAGCAGCTTCCCGGGGCGAGGGCTGCGGACCTTGACGCCCGCATGAGCTTCC 2359
Qy 2341 CTTCCATGGGCTCTGTGCGGCGCGGAGGCTCCCGGAGCAGCAGCAGCCCTGCTCCACAG 2400
Dh 2360 CTTCCATGGGCTCTGTGCGGCGCGGAGGCTCCCGGAGCAGCAGCAGCCCTGCTCCACAG 2419
Qy 2401 CGCCCAAGTCCCATGACACAGCAAGAGGCTGAGAAGTGGCGGCGCACGCGGAGCTGG 2460
Dh 2420 CGCCCAAGTCCCATGACACAGCAAGAGGCTGAGAAGTGGCGGCGCACGCGGAGCTGG 2479

QY	2461	CAGCAGCTACCCCTGCAGCCCTGTGCGGATCTCACTGGGTGAAGCCAGCTGGGCTCT	2520
Db	2480	CAGCAGCTACCCCTGCAGCCCTGTGCGGATCTCACTGGGTGAAGCCAGCTGGGCTCT	2539
QY	2521	GAGCTGGTGGAGACTGGAGAACCTTATCTAGCTCAGAGGATCGTAATACACCAAT	2580
Db	2540	GAGCTGGTGGAGACTGGAGAACCTTATCTAGCTCAGAGGATCGTAATACACCAAT	2599
QY	2581	CAGCACCCTGTGTCTAGCTCAGGGTCTGTGAATCCACCACCAATCCACTCTGTATCTAGCT	2640
Db	2600	CAGCACCCTGTGTCTAGCTCAGGGTCTGTGAATCCACCACCAATCCACTCTGTATCTAGCT	2659
QY	2641	ACTCTGATGGGGCTTGGAGACCTTTATGTCTTAGCTCAGGAGTTGTAAATACACCAATC	2700
Db	2660	ACTCTGATGGGGCTTGGAGAACCTTTATGTCTTAGCTCAGGAGTTGTAAATACACCAATC	2719
QY	2701	GGCACTGTATCTAGCTCAGAGTTTGTAAACACACCAATACAGACCTGTGTCTAGCTC	2760
Db	2720	GGCACTGTATCTAGCTCAGAGTTTGTAAACACACCAATACAGACCTGTGTCTAGCTC	2779
QY	2761	AGGGTATGTGAATGCACCAATCGACAGTCTGTATCTGGTACTTTCATGGGCATCCGTGT	2820
Db	2780	AGGGTATGTGAATGCACCAATCGACAGTCTGTATCTGGTACTTTCATGGGCATCCGTGT	2839
QY	2821	GAAGAAGACCAACCAACAGGCTTTGTGTGAGCAATAAAGCTTCTATCACCTGGGTGCAGGT	2880
Db	2840	GAAGAAGACCAACCAACAGGCTTTGTGTGAGCAATAAAGCTTCTATCACCTGGGTGCAGGT	2899
QY	2881	GGGCTGAGTCCGAAGAAGAGAGTCAAGCGAAGGAGATAAAGGTGGGGCCGTTTATAGGAT	2940
Db	2900	GGGCTGAGTCCGAAGAAGAGAGTCAAGCGAAGGAGATAAAGGTGGGGCCGTTTATAGGAT	2959
QY	2941	TTGGGTAGTAAAGAAAATTACAGTCAAAAGGGGCTTGTCTCTGGGGGACAGAGTGG	3000
Db	2960	TTGGGTAGTAAAGAAAATTACAGTCAAAAGGGGCTTGTCTCTGGGGGACAGAGTGG	3019
QY	3001	GGGGTCCGAAGTGTCTCAGTGGGGGCTTTTGTAGCCAGATAGCCAGAGAAAAGACT	3060
Db	3020	GGGGTCCGAAGTGTCTCAGTGGGGGCTTTTGTAGCCAGATAGCCAGAGAAAAGACT	3079
QY	3061	TTCAACAAGGTAAATGTCAATCTTAAAGGAAGACCCGCAATTCACCTCTTGTGGTG	3120
Db	3080	TTCAACAAGGTAAATGTCAATCTTAAAGGAAGACCCGCAATTCACCTCTTGTGTGGTG	3139
QY	3121	GAATGTCAATCAGTTAAAGTTGGGGCAGGGCATATTCATCTCTTGTGTATCTTCAGTTAC	3180
Db	3140	GAATGTCAATCAGTTAAAGTTGGGGCAGGGCATATTCATCTCTTGTGTATCTTCAGTTAC	3199
QY	3181	TTCAAGGCATCTGGGGCTATATGTGCAAGTTACAGGGATCCGATGGCTTGGGCT	3240
Db	3200	TTCAAGGCATCTGGGGCTATATGTGCAAGTTACAGGGATCCGATGGCTTGGGCT	3259
QY	3241	CAGAGGCTTGAAGCTACTCTGGTGGGGCCCTGGAGATGTTGTGTGAACACTGTAT	3300
Db	3260	CAGAGGCTTGAAGCTACTCTGGTGGGGCCCTGGAGATGTTGTGTGAACACTGTAT	3319
QY	3301	CTAGTTAATCTAGTGGGGCAGTGGAGAACTTTGTGTCTAGCTCAGGATGTAAACGCA	3360
Db	3320	CTAGTTAATCTAGTGGGGCAGTGGAGAACTTTGTGTCTAGCTCAGGATGTAAACGCA	3379
QY	3361	CCAATCAGCCCTGTCAAACACAGACACTGGGCTTACCAATCAGCAGGATGTGGGTGG	3420
Db	3380	CCAATCAGCCCTGTCAAACACAGACACTGGGCTTACCAATCAGCAGGATGTGGGTGG	3439
QY	3421	GGCAGATAAAGATTAAGCAGGCTCCGAGACGACAGTGGCAACGGCAGCAGGTCC	3480
Db	3440	GGCAGATAAAGATTAAGCAGGCTCCGAGACGACAGTGGCAACGGCAGCAGGTCC	3499
QY	3481	CTATTCACAATATGGCAGCTTGTCTTTGCTTTGGCGAATTAATCTGTACTGTCTCG	3540
Db	3500	CTATTCACAATATGGCAGCTTGTCTTTGCTTTGGCGAATTAATCTGTACTGTCTCG	3559
QY	3541	CTTTTGGGTCACACTGCTTTTATAGCTGTAAACACTCACACGAAAGTCTGCAGCTTC	3600

Db	3560	CTTTTGGGTCACACCTGCTTTTATGAGCTGTAACTCACACGAAAGTGCTGAGCTTC	3619
OY	3601	ACTCCTGAAGGCACCTAAGACCCACGACGCGGAGGAATGAACAATCCGGCGCGCT	3660
Db	3620	ACTCTGAAGGCACCTAAGACCCACGACGCGGAGGAATGAACAATCCGGCGCGCT	3679
OY	3661	GCTTTAAGAGCTATPAACCTCACCGCAGAGGTCGACGCTTCACTCCTCAGCCAGGAGA	3720
Db	3680	GCCTTAAGAGCTATPAACCTCACCGCAGAGGTCGACGCTTCACTCCTCAGCCAGGAGA	3739
OY	3721	CCAGGAAACCCACGAGAGGAAGAACTGCCAACATCTGAACATTCAGAAAGAACAACT	3780
Db	3740	CCAGGAAACCCACGAGAGGAAGAACTGCCAACATCTGAACATTCAGAAAGAACAACT	3799
OY	3781	CCAGATGCACACCTTAAGAGCTGTAACTACATCAGTCGAGAGGTCGCGGCTTCCTTTG	3840
Db	3800	CCAGATGCACACCTTAAGAGCTGTAACTACATCAGTCGAGAGGTCGCGGCTTCCTTTG	3859
OY	3841	AAGTCAGTGAGACCAAGCAGCTACAGTTTCGACACAGCCGAGAGTTTGAGATACG	3900
Db	3860	AAGTCAGTGAGACCAAGCAGCTACAGTTTCGACACAGCCGAGAGTTTGAGATACG	3919
OY	3901	CTGGGCAACATGATATGAATGGCCCTCTCGCAAAAAAAATTTCAAAAATTTGGCGG	3960
Db	3920	CTGGGCAACATGATATGAATGGCCCTCTCGCAAAAAAAATTTCAAAAATTTGGCGG	3979
OY	3961	AGCAATGATGTCCTCGCTCTGTCGCCACGTCAGCGGGGAGGCTAAAGTGAGAGATGCT	4020
Db	3980	AGCAATGATGTCCTCGCTCTGTCGCCACGTCAGCGGGGAGGCTAAAGTGAGAGATGCT	4039
OY	4021	TGAGCCTGGAGGTGAGACGTCAGTGAGCTGTGATTTGACACAGCCCTTAGCGTGG	4080
Db	4040	TGAGCCTGGAGGTGAGACGTCAGTGAGCTGTGATTTGACACAGCCCTTAGCGTGG	4099
OY	4081	GGACGAACTGAGACCCCTTTCCCTCCGGCAAAAAAATTTGCAAAAGTGAATTAAGGT	4140
Db	4100	GGACGAACTGAGACCCCTTTCCCTCCGGCAAAAAAATTTGCAAAAGTGAATTAAGGT	4159
OY	4141	GCCTATATMGCTAGCGCAGTCAGTCATGCCTGTAAATCCAGCACTTTGGGAACCGAG	4200
Db	4160	GCCTATATMGCTAGCGCAGTCAGTCATGCCTGTAAATCCAGCACTTTGGGAACCGAG	4219
OY	4201	GCGGCGGGGTCACTTAAGGTGAGAGTGTGAGACCGCTGGGCCCAACTGAGAGAAAGCC	4260
Db	4220	GCGGCGGGGTCACTTAAGGTGAGAGTGTGAGACCGCTGGGCCCAACTGAGAGAAAGCC	4279
OY	4261	ATCTCTTTCTAAAAATACAAATTTAGCCGGCTGTGGGGCACTGTGTGAGAGCATGGCTGTAA	4320
Db	4280	ATCTCTTTCTAAAAATACAAATTTAGCCGGCTGTGGGGCACTGTGTGAGAGCATGGCTGTAA	4339
OY	4321	TCCCACTACTCAGGAGGCTGAGGAGAGATACTTGAACCCAGAGAGGGGGGTTC	4380
Db	4340	TCCCACTACTCAGGAGGCTGAGGAGAGATACTTGAACCCAGAGAGGGGGGTTC	4399
OY	4381	AGTGAGCCGAGATCGTGCATTTGCACTCCACCACTCCAGCTGGGGCAACAAGGCCAA	4440
Db	4400	AGTGAGCCGAGATCGTGCATTTGCACTCCACCACTCCAGCTGGGGCAACAAGGCCAA	4459
OY	4441	CTCTGTCTTTAAAAAATAAAAAATGTCCTGACATTTAAGAGGTGTGCATATCAATG	4500
Db	4460	CTCTGTCTTTAAAAAATAAAAAATGTCCTGACATTTAAGAGGTGTGCATATCAATG	4519
OY	4501	TTGGCAGGCAACATTTTAAAGATGTGAGGTCGAGCTTCATGAGTCTCTTTAAAAAC	4560
Db	4520	TTGGCAGGCAACATTTTAAAGATGTGAGGTCGAGCTTCATGAGTCTCTTTAAAAAC	4579
OY	4561	CACCTTCAAGGCCAGGTGACGTGGCTATGCTATTAATCCAGCACTTTGGGAGCCGAG	4620
Db	4580	CACCTTCAAGGCCAGGTGACGTGGCTATGCTATTAATCCAGCACTTTGGGAGCCGAG	4639
OY	4621	GCGGATGATCACTGAGGTGAGAGTGTGAGACCGCTGACCAACCAACTGTGTAAAT	4680

D 4640 GCGGGTGGATCACCCTGAGGTCCAGAGCTTCAGACACCTGACCACCAACATGGTGAAT 4639
Q 4681 CCCACCTCTACTAAAAATACAAAAATTAGATGAGCAGTGGTGCATCCTGTAAATCCAC 4740
D 4700 CCCACCTCTACTAAAAATACAAAAATTAGATGAGCAGTGGTGCATCCTGTAAATCCAC 4759
Q 4741 CTACTTGGGAGGCTGAGGAGGAAATCATTAGAACAGGAGGCGGAGGTTGTAGAG 4800
D 4760 CTACTTGGGAGGCTGAGGAGGAAATCATTAGAACAGGAGGCGGAGGTTGTAGAG 4819
Q 4801 CCGAGATCGTGCATTGCTCCAGCTGAGCATGAGCAAAATCCATCTCAAAAAAAC 4860
D 4820 CCGAGATCGTGCATTGCTCCAGCTGAGCATGAGCAAAATCCATCTCAAAAAAAC 4879
Q 4861 AACAAACAAAAACCACTCTCTACTCCAGGAGAGCTGGGTACAGAGCTGGGCCATCAGT 4920
D 4880 AACAAACAAAAACCACTCTCTACTCCAGGAGAGCTGGGTACAGAGCTGGGCCATCAGT 4939
Q 4921 GCAAGGTGTGAGGCACAGAGCTAAAGCGGAGCTGCAAGACCGGGGACCAAGTAAACATG 4980
D 4940 GCAAGGTGTGAGGCACAGAGCTAAAGCGGAGCTGCAAGACCGGGGACCAAGTAAACATG 4999
Q 4981 TGTGAGATCAGTGTGAGATCAGACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040
D 5000 TGTGAGATCAGTGTGAGATCAGACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5059
Q 5041 AGCAACCAAGATGCCCCATCCAGTACCAACATCCATCTTCATCCAGAGATGCTGTTT 5100
D 5060 AGCAACCAAGATGCCCCATCCAGTACCAACATCCATCTTCATCCAGAGATGCTGTTT 5119
Q 5101 CTGGGACGCTGGGGGTAAATTAGAGACAGAGGTGACGCTCTGGGTGCTGGGTGCTGCTGCTG 5160
D 5120 CTGGGACGCTGGGGGTAAATTAGAGACAGAGGTGACGCTCTGGGTGCTGGGTGCTGCTGCTG 5179
Q 5161 TGGCCCAAGGAGGCTGTGAGGCTGTGAGAAACCTTCAGGCTTAGGCGGGGACGAGTGGC 5220
D 5180 TGGCCCAAGGAGGCTGTGAGGCTGTGAGAAACCTTCAGGCTTAGGCGGGGACGAGTGGC 5239
Q 5221 TCACGCTGTAAATCCAGCACTTTGGAGGCGGAGGCGGTGCATCAGAGTCAAGAGA 5280
D 5240 TCACGCTGTAAATCCAGCACTTTGGAGGCGGAGGCGGTGCATCAGAGTCAAGAGA 5299
Q 5281 TCGTGACCATCTGAGCTTAACAGGCTGAACCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5340
D 5300 TCGTGACCATCTGAGCTTAACAGGCTGAACCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5359
Q 5341 CCGGGGATGCTGGGCGGACCTGAGTTCAGCTACTCGGGAGGCTGAGGACAGAGAAATG 5400
D 5360 CCGGGGATGCTGGGCGGACCTGAGTTCAGCTACTCGGGAGGCTGAGGACAGAGAAATG 5419
Q 5401 GCGTGAAACCCGAGAGGACAGATTGTCAGTGACCGGAGATGCGGCACCTGCATCCAGCT 5460
D 5420 GCGTGAAACCCGAGAGGACAGATTGTCAGTGACCGGAGATGCGGCACCTGCATCCAGCT 5479
Q 5461 GGGCGACAGAGCAAGACTCCATCTGGAAGAAAGAAAAACGTTCAAGGCTGAGCCAGA 5520
D 5480 GGGCGACAGAGCAAGACTCCATCTGGAAGAAAGAAAAACGTTCAAGGCTGAGCCAGA 5539
Q 5521 GGGCGAGGCTGTAATCTGTCACTTACATGATGCTGGGCAAGGACATCTCTCCCTGGC 5580
D 5540 GGGCGAGGCTGTAATCTGTCACTTACATGATGCTGGGCAAGGACATCTCTCCCTGGC 5599
Q 5581 CCAAGTTCAGGAGGTTGGAATCAGATCCAGAGGCTCCCTTCAGATTAAACGCTGATGGTT 5640
D 5600 CCAAGTTCAGGAGGTTGGAATCAGATCCAGAGGCTCCCTTCAGATTAAACGCTGATGGTT 5659
Q 5641 TAAGATGAGAAATGGGGAGTTTCCCTCTCTACCCCAAGCCCGGTGTCCACTTCAAGGT 5700
D 5660 TAAGATGAGAAATGGGGAGTTTCCCTCTCTACCCCAAGCCCGGTGTCCACTTCAAGGT 5719
Q 5701 GAATGACAGAGGAAATGACGTGTCCAAATCCGCAAGTTCGCAAGGCTTGGGAGCCCTAC 5760
D 5720 GAATGACAGAGGAAATGACGTGTCCAAATCCGCAAGTTCGCAAGGCTTGGGAGCCCTAC 5779

Q 5761 TGTGAGGCTGTGACGAGAGAGTGAAGGTTCAGGTGAGCCAAATCCCTCGAAGGCTCTG 5820
D 5780 TGTGAGGCTGTGACGAGAGAGTGAAGGTTCAGGTGAGCCAAATCCCTCGAAGGCTCTG 5839
Q 5821 CCTCATTCGGGACAGACATCCGGTTTCTCTGCTCTACCGGGATTTAGGGGCTTACG 5880
D 5840 CCTCATTCGGGACAGACATCCGGTTTCTCTGCTCTACCGGGATTTAGGGGCTTACG 5899
Q 5881 CGAATGATGATGGGGGCGGGGGGCTTGTGGGGAGATTCCCAAGTAAATCACTTGGGA 5940
D 5900 CGAATGATGATGGGGGCGGGGGGCTTGTGGGGAGATTCCCAAGTAAATCACTTGGGA 5959
Q 5941 CAGGACAGCTCGAATCTTTCAGTGTCTCTATCCAAATGTGGGTGGGACAGAGCCAA 6000
D 5960 CAGGACAGCTCGAATCTTTCAGTGTCTCTATCCAAATGTGGGTGGGACAGAGCCAA 6019
Q 6001 GAACCAATGTCCTTATCTCAGATGAGGCTGAGGAGTTCACAGAGTTCACAGAGGCTCCG 6060
D 6020 GAACCAATGTCCTTATCTCAGATGAGGCTGAGGAGTTCACAGAGTTCACAGAGGCTCCG 6079
Q 6061 AGAATTTGGGGGTAGGAATGGGAGCAACAGGCTTCTTTTCTCTTCTTCTTCTTCTTCTT 6120
D 6080 AGAATTTGGGGGTAGGAATGGGAGCAACAGGCTTCTTTTCTCTTCTTCTTCTTCTTCTT 6139
Q 6121 GCGTTGGGGGACAGCTTTCAGAAATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 6180
D 6140 GCGTTGGGGGACAGCTTTCAGAAATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 6199
Q 6181 TCTGCCACAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6235
D 6200 TCTGCCACAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6254

RESULT 4
US-09-525-160B-5
; Sequence 5, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evgenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525.160B
; CURRENT FILING DATE: 2000-03-14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-5

Query Match 100.0%; Score 6235; DB 4; Length 6679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 GATCACTTGGAGACAGTACTTCAAGACCACTGAGGACAGATAGGAGAGTGTCTACG 60
D 20 GATCACTTGGAGACAGTACTTCAAGACCACTGAGGACAGATAGGAGAGTGTCTACG 79
Q 61 AAAAATCAAAAATTTATGGCGGCGATGTGGCTCACGTCTGTAAATCCCTGAACCTTGGG 120
D 80 AAAAATCAAAAATTTATGGCGGCGATGTGGCTCACGTCTGTAAATCCCTGAACCTTGGG 139
Q 121 ACATCAAGGAGAGTGAATCTGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
D 140 ACATCAAGGAGAGTGAATCTGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 199
Q 181 GAAACCTTATCTCCACTTAAAAATACAAAAATTTAGCCAGGACAGTGTGGCAGCACTGTGA 240
D 200 GAAACCTTATCTCCACTTAAAAATTTAGCCAGGACAGTGTGGCAGCACTGTGA 259
Q 241 ATCCGCGCTACTCAGAGGCTGAGGACAGAGAAATCACTTGAACCCAGAGAGGCGAGGTTG 300

Db 260 ATCCGGCTACTCAGAGAGCTGAGGAGAGATACATTGAAACCCAGAGAGGGAGTTG 319
Qy 301 CAGTGAAGCTGAGATCACACACTGCATCCAGCCCTGGGTGACAGAGCAAGACTATCTC 360
Db 320 CAGTGAAGCTGAGATCACACACTGCATCCAGCCCTGGGTGACAGAGCAAGACTATCTC 379
Qy 361 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
Db 380 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 439
Qy 421 CTACTCAGAGAGCTGAGGTGGAGAGATCACTTGAACCTGGGGCACTCAAGCTAGAGTA 480
Db 440 CTACTCAGAGAGCTGAGGTGGAGAGATCACTTGAACCTGGGGCACTCAAGCTAGAGTA 499
Qy 481 GCCAAGATCATGCCCTACACTCCAGCCTGGGCAACAGAGAGAGACCTGTCTAAANA 540
Db 500 GCCAAGATCATGCCCTACACTCCAGCCTGGGCAACAGAGAGAGACCTGTCTAAANA 559
Qy 541 AATAATAATAATAAAGAAAAACAGCTGTGTTATGTCTCTGCTGCATCATCATCTACT 600
Db 560 AATAATAATAATAAAGAAAAACAGCTGTGTTATGTCTCTGCTGCATCATCATCTACT 619
Qy 601 ATGTATATAGTTTGCACAACTCAAGATCCAGATAGTCAATTTTAAAGCCTGTGGCCGT 660
Db 620 ATGTATATAGTTTGCACAACTCAAGATCCAGATAGTCAATTTTAAAGCCTGTGGCCGT 679
Qy 661 ATGGCTCTGTACATCACTCTGCTGCTCTCTTCTAGACAAAGACGCTAATAACAT 720
Db 680 ATGGCTCTGTACATCACTCTGCTGCTCTCTTCTAGACAAAGACGCTAATAACAT 739
Qy 721 ACATACATGAATTTTATAGACATCGAGATTGTAATTCATATGATTTTACATTTAT 780
Db 740 ACATACATGAATTTTATAGACATCGAGATTGTAATTCATATGATTTTACATTTAT 799
Qy 781 AAAATTAATCTTTTAAAAATTTTCCCTTACCATTAAAAAGTAAAAAGCCGGCAGGCG 840
Db 800 AAAATTAATCTTTTAAAAATTTTCCCTTACCATTAAAAAGTAAAAAGCCGGCAGGCG 859
Qy 841 GCCATCTGCACGCCCTGTAAATTCACAGACTTTGGGAGGCTGAGGTGGGCAAGTCTGAG 900
Db 860 GCCATCTGCACGCCCTGTAAATTCACAGACTTTGGGAGGCTGAGGTGGGCAAGTCTGAG 919
Qy 901 ATCAACAGTTGAGAGACAGCCTGGCCACATAGCAAAACCCATTCTACTAAAAATAA 960
Db 920 ATCAACAGTTGAGAGACAGCCTGGCCACATAGCAAAACCCATTCTACTAAAAATAA 979
Qy 961 AAAATTAAGCTGGGCTAGTGTGTCACACCTGTGATCCAGCTACTTGGGAGGCTGAGCA 1020
Db 980 AAAATTAAGCTGGGCTAGTGTGTCACACCTGTGATCCAGCTACTTGGGAGGCTGAGCA 1039
Qy 1021 GGAGAAATGCTTGAACCTGGGAGAGGGAGTTGCAAGTGAAGCCACATATGCACTGCAC 1080
Db 1040 GGAGAAATGCTTGAACCTGGGAGAGGGAGTTGCAAGTGAAGCCACATATGCACTGCAC 1099
Qy 1081 TCCAGCCTGGGTGACAGAGTGAAGTGTCTCAACGAAAAAAGGTAAGCCAT 1140
Db 1100 TCCAGCCTGGGTGACAGAGTGAAGTGTCTCAACGAAAAAAGGTAAGCCAT 1159
Qy 1141 TCTTAATTCAGTGTACATCAGTGTACATCACTAGGTCTGCGTACTCTGTGAGGAT 1200
Db 1160 TCTTAATTCAGTGTACATCAGTGTACATCACTAGGTCTGCGTACTCTGTGAGGAT 1219
Qy 1201 ACCTGAGAGTGAAGTGTGCTGTCACAGACATACAAATTCATTCATTAAGTACACT 1260
Db 1220 ACCTGAGAGTGAAGTGTGCTGTCACAGACATACAAATTCATTCATTAAGTACACT 1279
Qy 1261 ACCAAGTGGCATCCAGAGAGTTTATTTTATACAAATTCATTCATTCATTCATTCAT 1320
Db 1280 ACCAAGTGGCATCCAGAGAGTTTATTTTATACAAATTCATTCATTCATTCATTCAT 1339
Qy 1321 GAGAGTACTCAGATCTTTTACAAAGATGCTTAAGCCAGTACAGATGAACAGAGA 1380

Db 1340 GAGAGTACTCAGATCTTTTACAAAGATGCTTAAGCCAGTACCAGATGAACAGAGA 1399
Qy 1381 AGTGGAGGGGGAAGCTGCCAGCCCTTCTAACCATGAAGAAATACCTGGTAGACCTTCT 1440
Db 1400 AGTGGAGGGGGAAGCTGCCAGCCCTTCTAACCATGAAGAAATACCTGGTAGACCTTCT 1459
Qy 1441 GGATGCTGGAAGGATGAATTAACGGGGGTCTGGAGCCCGCCCTGTACAGTCACTGTG 1500
Db 1460 GGATGCTGGAAGGATGAATTAACGGGGGTCTGGAGCCCGCCCTGTACAGTCACTGTG 1519
Qy 1501 ACTTCTGAGCCCTCCAGTCCAGTCCAGCCCATGTGTATGAGCCAGTAAATGAGCCCT 1560
Db 1520 ACTTCTGAGCCCTCCAGTCCAGTCCAGCCCATGTGTATGAGCCAGTAAATGAGCCCT 1579
Qy 1561 CACTCTGTGTGTGTCTTATATCTCCCATGTGGGGCTGAAGTCTGATGAGCCCTAT 1620
Db 1580 CACTCTGTGTGTGTCTTATATCTCCCATGTGGGGCTGAAGTCTGATGAGCCCTAT 1639
Qy 1621 TCAAGATGTACAGCTTTCTTGACAGAAAGTGTGTACAGAAACAGCAGGGGCTTGCA 1680
Db 1640 TCAAGATGTACAGCTTTCTTGACAGAAAGTGTGTACAGAAACAGCAGGGGCTTGCA 1699
Qy 1681 AGATGATCTAAGTCAAAATCTTACCTGGCTCAGCCACAGCAGTATGTGTATCTGAAC 1740
Db 1700 AGATGATCTAAGTCAAAATCTTACCTGGCTCAGCCACAGCAGTATGTGTATCTGAAC 1759
Qy 1741 AAGTTTTTCACTTCTTGAGGCCATTCCTTGGCTTACAAACACAGTGTGTGACAGGA 1800
Db 1760 AAGTTTTTCACTTCTTGAGGCCATTCCTTGGCTTACAAACACAGTGTGTGACAGGA 1819
Qy 1801 TGAATGAGGAAGTCCCTTACACCGTAAATCCACACTTTGGGAGGCCAAGGGGGTGG 1860
Db 1820 TGAATGAGGAAGTCCCTTACACCGTAAATCCACACTTTGGGAGGCCAAGGGGGTGG 1879
Qy 1861 ATGGCTGAGCCTGAGAGGTGACAGATCCCGGACAGTCTCTCACAGCCCTGTCTCTC 1920
Db 1880 ATGGCTGAGCCTGAGAGGTGACAGATCCCGGACAGTCTCTCTCACAGCCCTGTCTCTC 1939
Qy 1921 GGGGCTCTCTGCTGGGCTGCCACTTGGGTGGGCACTTGAAGAGCCCTTACAGCCACCG 1980
Db 1940 GGGGCTCTCTGCTGGGCTGCCACTTGGGTGGGCACTTGAAGAGCCCTTACAGCCACCG 1999
Qy 1981 CTGCACTGTGGGAGCCCTTCTTGGGCTGGGCAAGGAGCAGGAGCCCTCTCACCTTGC 2040
Db 2000 CTGCACTGTGGGAGCCCTTCTTGGGCTGGGCAAGGAGCAGGAGCCCTCTCACCTTGC 2059
Qy 2041 AGGAGGTGTGAGGAGAGAGCTCAAGCAGAGAACGGGGCTGCGACAGGCGCTGCGGGC 2100
Db 2060 AGGAGGTGTGAGGAGAGAGCTCAAGCAGAGAACGGGGCTGCGACAGGCGCTGCGGGC 2119
Qy 2101 CAGCTGGAATTCGGGGTGGGCTGGGCTTGGGGGCCCCGCACTGAGAGCAAGGCGCAG 2160
Db 2120 CAGCTGGAATTCGGGGTGGGCTGGGCTTGGGGGCCCCGCACTGAGAGCAAGGCGCAG 2179
Qy 2161 CCCCTCAGAGCCCGGGGCAATGAGAGGCTTACACCCGGGCGCAGGCGCTGCGGAGGTGT 2220
Db 2180 CCCCTCAGAGCCCGGGGCAATGAGAGGCTTACACCCGGGCGCAGGCGCTGCGGAGGTGT 2239
Qy 2221 ACTGGGTGCCAGCAGTGCAGCCCGCGGCGTGTGCTGCTGATTTTCTACTGGGC 2280
Db 2240 ACTGGGTGCCAGCAGTGCAGCCCGCGGCGTGTGCTGCTGATTTTCTACTGGGC 2299
Qy 2281 CTTAGCAGCTTCCCGGCGGAGGCTTGGGAGCTTGCAGCCCGCATGCTGAGCCCTTC 2340
Db 2300 CTTAGCAGCTTCCCGGCGGAGGCTTGGGAGCTTGCAGCCCGCATGCTGAGCCCTTC 2359
Qy 2341 CTTCCATGGGCTCTGTGGGCGGAGGCTCCCGCAGCAGCAACCCCTGCTCCACAG 2400
Db 2360 CTTCCATGGGCTCTGTGGGCGGAGGCTCCCGCAGCAGCAACCCCTGCTCCACAG 2419
Qy 2401 CGCCAGTCCCATGCACACAGCAAGGCTGAGAACTGGGGGCGCACGGCACGGGAGTGG 2460
Db 2420 CGCCAGTCCCATGCACACAGCAAGGCTGAGAACTGGGGGCGCACGGGAGTGG 2479

QY	2461	CAGCAGCTACCCCTGCAGCCCTGCTGGCGAATCCACTGGGTGAAGCCAGCTGGGCTCT	2520
Db	2480	CAGCAGCTACCCCTGCAGCCCTGCTGGCGAATCCACTGGGTGAAGCCAGCTGGGCTCT	2539
QY	2521	GAGCTGGTGGAGACTGGAGAACCTTATCTACGTCAGGGGATCGTAATATACCAAT	2580
Db	2540	GAGCTGGTGGAGACTGGAGAACCTTATCTACGTCAGGGGATCGTAATATACCAAT	2599
QY	2581	CAGCAGCCCTGTGTAGCTCAGGGGTCTGTGAATGCACCAATCCACTCTGTATCTAGCT	2640
Db	2600	CAGCAGCCCTGTGTAGCTCAGGGGTCTGTGAATGCACCAATCCACTCTGTATCTAGCT	2659
QY	2641	ACTCTGATGGGGCCTTGGAGAACCTTTATGTCTTAGCTCAGGGATTGTAAATACCAATC	2700
Db	2660	ACTCTGATGGGGCCTTGGAGAACCTTTATGTCTTAGCTCAGGGATTGTAAATACCAATC	2719
QY	2701	GGCACTCTGTATCTAGCTCAAGGTTTGTAAACACACCAATATAGCAACCCCTGTCTAGTC	2760
Db	2720	GGCACTCTGTATCTAGCTCAAGGTTTGTAAACACCAATATAGCAACCCCTGTCTAGTC	2779
QY	2761	AGGCTATGTGAATGCACCAATCGACAGTCTGTATCTGGCTACTTTCATGGGCATCCGTGT	2820
Db	2780	AGGCTATGTGAATGCACCAATCGACAGTCTGTATCTGGCTACTTTCATGGGCATCCGTGT	2839
QY	2821	GAAAGAGACACCAAAACAGGCTTTGTGTAGCAATAAAGCTTCATACACCTGGGGGACAGT	2880
Db	2840	GAAAGAGACACCAAAACAGGCTTTGTGTAGCAATAAAGCTTCATACACCTGGGGGACAGT	2899
QY	2881	GGGCTGAGTCCGAAAAAGAGTCAAGCGAAGGAGATAAGGGTGGGGCCGTTTATAGAT	2940
Db	2900	GGGCTGAGTCCGAAAAAGAGTCAAGCGAAGGAGATAAGGGTGGGGCCGTTTATAGAT	2959
QY	2941	TTGGGTAGGTAAAGAAAAATTACAGTCAAAAGGGGTTTGTCTCTGGGGGACAGAGTGG	3000
Db	2960	TTGGGTAGGTAAAGAAAAATTACAGTCAAAAGGGGTTTGTCTCTGGGGGACAGAGTGG	3019
QY	3001	GGGGTCGCAAGGTGCTCAGTGGGGGTGCTTTTGAAGCAGAGATAGCCAGAGAAAAAGACT	3060
Db	3020	GGGGTCGCAAGGTGCTCAGTGGGGGTGCTTTTGAAGCAGAGATAGCCAGAGAAAAAGACT	3079
QY	3061	TTCAACAAGTATGTCAATTAAGGCAGAGACCCGGCATTTACACCTCTTTGTGGTG	3120
Db	3080	TTCAACAAGTATGTCAATTAAGGCAGAGACCCGGCATTTACACCTCTTTGTGGTG	3139
QY	3121	GAATGTCAATCAAGTTAAGTTGGGGCAGGGCATATTCACCTCTTTTGTGATTTTCAATTAC	3180
Db	3140	GAATGTCAATCAAGTTAAGTTGGGGCAGGGCATATTCACCTCTTTTGTGATTTTCAATTAC	3199
QY	3181	TTCAAGGCCATCTGGGGCGTATATGTGCAAGTTACAGGGGATCGAATGGCTTGGGCT	3240
Db	3200	TTCAAGGCCATCTGGGGCGTATATGTGCAAGTTACAGGGGATCGAATGGCTTGGGCT	3259
QY	3241	CAGAGGCTTGACAGCTACTCTGTGGTGGGGCCTTGGAGAAATGTTTGTGTGCACACTCTGAT	3300
Db	3260	CAGAGGCTTGACAGCTACTCTGTGGTGGGGCCTTGGAGAAATGTTTGTGTGCACACTCTGAT	3319
QY	3301	CTAGTTAATCAAGTGGGGAGCTGGAGAACCTTTGTGCTTACTCAGGGATTGTAAAGCA	3360
Db	3320	CTAGTTAATCAAGTGGGGAGCTGGAGAACCTTTGTGCTTACTCAGGGATTGTAAAGCA	3379
QY	3361	CCAATTCAGGCGCTCTGTAACAAACAGACACTGGGCTTACCAATCAGCAGATGTGGGTGG	3420
Db	3380	CCAATTCAGGCGCTCTGTAACAAACAGACACTGGGCTTACCAATCAGCAGATGTGGGTGG	3439
QY	3421	GGCCAGATATAGAGATAAAGCAGAGCTGCCCGAGCCAGCAGTGGCCAACGCCACAGGCTC	3480
Db	3440	GGCCAGATATATAGAGATAAAGCAGAGCTGCCCGAGCCAGCAGTGGCCAACGCCACAGGCTC	3499
QY	3481	CTATCCCAAAATATGGCAGCTTGTCTCTTTTGTCTGTGGCATTAATCTTGTACTGCTCG	3540
Db	3500	CTATCCCAAAATATGGCAGCTTGTCTCTTTTGTCTGTGGCATTAATCTTGTACTGCTCG	3559

QY	3541	CTTTTGGGTCACACTGCTTTTATATAGCTGTAACTCAACACGAAGCTCTGACGTTT	3600
Db	3560	CTTTTGGGTCACACTGCTTTTATATAGCGTAACTCAACCAAGGCTCTGACGTTT	3619
QY	3601	ACTCGAAGCCACTAAGACCAAGACCCACCGGGGGAATGAACAACCTCGGGCCGGCT	3660
Db	3620	ACTCGTGAAGCCACTAAGACCAAGACCCACCGGGGGAATGAACAACCTCGGGCCGGCT	3679
QY	3661	GCTTAAAGCTATAACACTCAACCGGAAGGTGTGACGTTCTACTCTCCAGCAGCAGA	3720
Db	3680	GCTTAAAGGCTATAACACTCAACCGGAAGGTGTGACGTTCTACTCTCCAGCAGCAGA	3739
QY	3721	CCAGCAACCCACGAGAGAAGAACTGGCAACATCTGACATCAAGAAAGAACAACT	3780
Db	3740	CCAGCAACCCACGAGAGAAGAACTGGCAACATCTGACATCAAGAAAGAACAACT	3799
QY	3781	CCAGATGCACACCTTAAAGCGTAAACATCACTCGAGGGGTCCGGGCTTCCTTG	3840
Db	3800	CCAGATGCACACCTTAAAGCGTAAACATCACTCGAGGGGTCCGGGCTTCCTTG	3859
QY	3841	AAGTAGTAGACCAAGCACTCACAGTTTCGACACAAAGCCACGAGATTGTAGATAGC	3900
Db	3860	AAGTAGTAGACCAAGCACTCACAGTTTCGACACAAAGCCACGAGATTGTAGATAGC	3919
QY	3901	CTGGGCACATGATGAAATGCCCTCTCTGCAAAAAAAAAAAAAAAAAATTATGCGCG	3960
Db	3920	CTGGGCACATGATGAAATGCCCTCTCTGCAAAAAAAAAAAAAAAAAATTATGCGCG	3979
QY	3961	AGCATGTGTGCTCGTACCTGTGTGCCAGCTACGCGGAGGCTAAAGTGGAGATGCT	4020
Db	3980	AGCATGTGTGCTCGTACCTGTGTGCCAGCTACGCGGAGGCTAAAGTGGAGATGCT	4039
QY	4021	TGAGCCTGGAGGTGAAGACTGCAGTAGCTGTGATTGTACCAAGCCCTAGCTGGG	4080
Db	4040	TGAGCCTGGAGGTGAAGACTGCAGTAGCTGTGATTGTACCAAGCCCTAGCTGGG	4099
QY	4081	GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAAAAAATTACAAAAGTGAATAAAGGT	4140
Db	4100	GGACAGACTGAGACCTGTTTCCCTCCGCAAAAAAAAAATTACAAAAGTGAATAAAGGT	4159
QY	4141	GGCTGATATGGGCTAAGGCGAGTGGCTATGCCCTGTATCCACAGACTTGGGAAGCGAG	4200
Db	4160	GGCTGATATGGGCTAAGGCGAGTGGCTATGCCCTGTATCCACAGACTTGGGAAGCGAG	4219
QY	4201	GGGGGGGGTCACTTAAGGTCAAGAGATGTGAGACCAAGCCCTGGCAACATGGAGAAACCC	4260
Db	4220	GGGGGGGGTCACTTAAGGTCAAGAGATGTGAGACCAAGCCCTGGCAACATGGAGAAACCC	4279
QY	4261	ATCTCTTCAAAAAATCAAAAAATTAGCCGCGTGTGGGGGCGAGTGGTGGAGATGCCGTAA	4320
Db	4280	ATCTCTTCAAAAAATCAAAAAATTAGCCGCGTGTGGGGGCGAGTGGTGGAGATGCCGTAA	4339
QY	4321	TCCCAAGCTACTAGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGGGTTGC	4380
Db	4340	TCCCAAGCTACTAGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGGGTTGC	4399
QY	4381	AGTAGGCCAGATTCGTCATTGCACTCCACCCACTTCACAGCTGGGCAACAAGACCAA	4440
Db	4400	AGTAGGCCAGATTCGTCATTGCACTCCACCCACTTCACAGCTGGGCAACAAGACCAA	4459
QY	4441	CTCTGCTTAAAAAAAAAAAAAAAAAGTGCCTGACATATTAAGAGTGTGCAATGCAATAG	4500
Db	4460	CTCTGCTTAAAAAAAAAAAAAAAAAGTGCCTGACATATTAAGAGTGTGCAATGCAATAG	4519
QY	4501	TTGCGCAGGCAACATGTTTAAGATGTGAGACTCTGCTTCCATGGTCTGTTAAAAAC	4560
Db	4520	TTGCGCAGGCAACATGTTTAAGATGTGAGACTCTGCTTCCATGGTCTGTTAAAAAC	4579
QY	4561	CACCCCAAGGGCAGGTGAGTGGCTCAATGCCATATATCCACACCTTTGGGAGGCGAG	4620
Db	4580	CACCCCAAGGGCAGGTGAGTGGCTCAATGCCATATATCCACACCTTTGGGAGGCGAG	4639
QY	4621	GGGGGTGATCACTTAGTCAAGAGATTGTGAGACCAAGCCCTGACCAACATGTTGTAAT	4680

Db 4640 GCGGGTGGATCAGCTGAGGCTGAGAGTTCAGAGCAGCAGCTGACCAACCATGGTGAAT 4639
QY 4681 CCCAGCTCTACTAAAAATACAAAAATTAGATGAGCATGGTGGTCATCCTGTAAATCCAC 4740
Db 4700 CCCAGCTCTACTAAAAATACAAAAATTAGATGAGCATGGTGGTCATCCTGTAAATCCAC 4739
QY 4741 CTACTTGGGAGGCTGAGGAGGAGAAATCATAAGAACAGGAGGCGGAGGTTGTAGTAC 4800
Db 4760 CTACTTGGGAGGCTGAGGAGGAGAAATCATAAGAACAGGAGGCGGAGGTTGTAGTAC 4819
QY 4801 CCAGATCGTGCCATTGCTCCTCCAGCTGAGCATGAGCGAAATCATGTCAAAAAAAC 4860
Db 4820 CCAGATCGTGCCATTGCTCCTCCAGCTGAGCATGAGCGAAATCATGTCAAAAAAAC 4879
QY 4861 AACACAAAAACCCACTCTCTACTCCAGGAGCTGGGTACAGAGTGGGCGACATCATGT 4920
Db 4880 AACACAAAAACCCACTCTCTACTCCAGGAGCTGGGTACAGAGTGGGCGACATCATGT 4939
QY 4921 GCAAGGTGCTGAGGCAACAGAGCTAAAGCGGAGCTGCGAGCAGCGGAGCAAGATTAACGTG 4980
Db 4940 GCAAGGTGCTGAGGCAACAGAGCTAAAGCGGAGCTGCGAGCAGCGGAGCAAGATTAACGTG 4999
QY 4981 TGTAGATCAGTGTGTAGATCAGAGCTCCCTGCCATTGGTGAACACAGGCGGCCCA 5040
Db 5000 TGTAGATCAGTGTGTAGATCAGAGCTCCCTGCCATTGGTGAACACAGGCGGCCCA 5059
QY 5041 AGACACAGAGATGGCCCCATCCAGTACACACATCCACTTCTCATCCAGAGATGTCTGTTT 5100
Db 5060 AGACACAGAGATGGCCCCATCCAGTACACACATCCACTTCTCATCCAGAGATGTCTGTTT 5119
QY 5101 CTGGCAGCGCTGGGGTAAATTAGAGACGAAGAGTGACAGTCTGGGTGGTGGTCACTGAC 5160
Db 5120 CTGGCAGCGCTGGGGTAAATTAGAGACGAAGAGTGACAGTCTGGGTGGTGGTCACTGAC 5179
QY 5161 TGCCCCAGGAGGCGCTTGTGGCTGTAGAAAAGTTCAGGCTTAGGCGGGGCAAGCTGCGC 5220
Db 5180 TGCCCCAGGAGGCGCTTGTGGCTGTAGAAAAGTTCAGGCTTAGGCGGGGCAAGCTGCGC 5239
QY 5221 TCACGCTGTAAATCCAGACATTTGGGAGGCGGAGGCGGGTGCATCCAGAGTCAAGAGA 5280
Db 5240 TCACGCTGTAAATCCAGACATTTGGGAGGCGGAGGCGGGTGCATCCAGAGTCAAGAGA 5299
QY 5281 TCGTGAACATCTGGCTTAACAGGTGAACCCCGTCTCTCTAAAAATACAAAAAATTGG 5340
Db 5300 TCGTGAACATCTGGCTTAACAGGTGAACCCCGTCTCTCTAAAAATACAAAAAATTGG 5339
QY 5341 CCGGGCATGTGGCGGGCACCTGTAGTTCAGCTACTCGGAGGCTTAGGCAAGAGATG 5400
Db 5360 CCGGGCATGTGGCGGGCACCTGTAGTTCAGCTACTCGGAGGCTTAGGCAAGAGATG 5419
QY 5401 GCGTGAACCCGAGAGGAGGAGTTCAGAGTACCGGAGATGCGCCACTGCACTCCAGCT 5460
Db 5420 GCGTGAACCCGAGAGGAGGAGTTCAGAGTACCGGAGATGCGCCACTGCACTCCAGCT 5479
QY 5461 GGGCGACAGAGCAAGACTCCATCTGAAAAAGAAAAAGTTCAGGTCTGAGCGAGA 5520
Db 5480 GGGCGACAGAGCAAGACTCCATCTGAAAAAGAAAAAGTTCAGGTCTGAGCGAGA 5539
QY 5521 GGGCGACAGGCTGTAAATCTGTCACTTACATGACCTTGGGCAAGGCACTTCTCCCTGGC 5580
Db 5540 GGGCGACAGGCTGTAAATCTGTCACTTACATGACCTTGGGCAAGGCACTTCTCCCTGGC 5599
QY 5581 CCAGTTCAGGAGGTTGGAATCGACTCAAGGTCCTTCCAGATTAACGTCGATGGTTTC 5640
Db 5600 CCAGTTCAGGAGGTTGGAATCGACTCAAGGTCCTTCCAGATTAACGTCGATGGTTTC 5659
QY 5641 TAAGATGAGAGATGGGAGGTTTCCCTCTCACCCGAGCCCGTTCACATTCAGAGT 5700
Db 5660 TAAGATGAGAGATGGGAGGTTTCCCTCTCTCACCCGAGCCCGTTCACATTCAGAGT 5719
QY 5701 GAATGACAGAGGAATCAGTGTCCCAATCCGCAAGTTCCAAGGCCCTTGGGACCTTAC 5760
|||||

Db 5720 GAATGACAGAGGAATCAGTGTCCCAATCCCGCAAGTTCACAAAGCCCTTGGGACCTTAC 5779
QY 5761 TGTAGGAGTGTGTGACAGAGAGGTGAAGGTTCAGGTGAGCCAAATCCCTCGAAGGCTTTG 5820
Db 5780 TGTAGGAGTGTGTGACAGAGAGGTGAAGGTTCAGGTGAGCCAAATCCCTCGAAGGCTTTG 5839
QY 5821 CCTCATTTGGGAGGAGATCCGGTTTCTGCTGTGCTTACCCGGGATTTAGAGGCTTTAGC 5880
Db 5840 CCTCATTTGGGAGGAGATCCGGTTTCTGCTGTGCTTACCCGGGATTTAGAGGCTTTAGC 5899
QY 5881 CGAATGAGTATGGGAGGAGGAGGAGGTTTCTGAGGAGTTCCAGTAAATCAACTTGGGA 5940
Db 5900 CGAATGAGTATGGGAGGAGGAGGAGGTTTCTGAGGAGTTCCAGTAAATCAACTTGGGA 5959
QY 5941 CAGGACAGCTGGAATTTGAGTGTGCTATCCAAAGTGTGGGTGGGACAGAGCCAA 6000
Db 5960 CAGGACAGCTGGAATTTGAGTGTGCTATCCAAAGTGTGGGTGGGACAGAGCCAA 6019
QY 6001 GACCCATGTCCTTATCTCAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6060
Db 6020 GACCCATGTCCTTATCTCAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6079
QY 6061 AGAGTTGGGAGTGAAGTGGAGCAACAGAGGCTTCTTTTCTCTTAAATTTGGG 6120
Db 6080 AGAGTTGGGAGTGAAGTGGAGCAACAGAGGCTTCTTTTCTCTTAAATTTGGG 6139
QY 6121 GCGTTGGGAGGAGGCTTGAAGTCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6180
Db 6140 GCGTTGGGAGGAGGCTTGAAGTCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6199
QY 6181 TCTGCCAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6235
Db 6200 TCTGCCAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6254

RESULT 5
US-09-305-384-6
: Sequence 6, Application US/09305384
: Patent No. 6242218
: GENERAL INFORMATION:
: APPLICANT: Treco, Douglas A.
: APPLICANT: Heartlein, Michael W.
: APPLICANT: Selden, Richard F.
: TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
: FILE REFERENCE: 07236/017001
: CURRENT APPLICATION NUMBER: US/09/305,384
: EARLIER FILING DATE: 1999-05-05
: EARLIER APPLICATION NUMBER: US 60/084,649
: EARLIER FILING DATE: 1998-05-07
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 2834
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-305-384-6

Query Match 45.5%; Score 2834; DB 3; Length 2834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1890 CCGGCAATCCGACAGGCGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1949
QY 1890 CCGGCAATCCGACAGGCGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1950
Db 1950 GGTGGCACTTGAAGAGGCGCTTCAAGCCAGCGTGAAGTGGAGAGCCCTTCTGGGCTG 2009
QY 1950 GGTGGCACTTGAAGAGGCGCTTCAAGCCAGCGTGAAGTGGAGAGCCCTTCTGGGCTG 2009
Db 61 GGTGGCACTTGAAGAGGCGCTTCAAGCCAGCGTGAAGTGGAGAGCCCTTCTGGGCTG 120
QY 2010 GCCAAGGCCAGAGCGGCTTCTCAAGCTTCAAGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGG 2069
Db 121 GCCAAGGCCAGAGCGGCTTCTCAAGCTTCAAGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 180
|||||

QY 2070 GGAACCGGGGCTGCGCAGCGCGCTTGCGGGGCCAGCTGGAAGTTCCGGGTGGCGTGGCTT 2129
|||||
Db 181 GGAACCGGGGCTGCGCAGCGCGCTTGCGGGGCCAGCTGGAAGTTCCGGGTGGCGTGGCTT 240
|||||
QY 2130 GGCAGGCGCCGCACTGCGAGCAGCGGGCCAGCCCTCCAGGCGCCCGGGCAATGAGAGCT 2189
|||||
Db 241 GGGGGGGCGCGCACTGCGAGCAGCGGGCCAGCCCTCCAGGCGCCCGGGCAATGAGAGCT 300
|||||
QY 2190 TACACCCGGGGCAGCGGCTGCGGAGGGTGTACTGGGTCCCGCAGCAGTGCACGCCGCC 2249
|||||
Db 301 TACACCCGGGGCAGCGGCTGCGGAGGGTGTACTGGGTCCCGCAGCAGTGCACGCCGCC 360
|||||
QY 2250 GGGGCTGTGCTGCTGATTTCTCACTGGGCGCTTACAGCCTTCCGCGGGGAGGGGCTC 2309
|||||
Db 361 GGGGCTGTGCTGCTGATTTCTCACTGGGCGCTTACAGCCTTCCGCGGGGAGGGGCTC 420
|||||
QY 2310 GGGACCTGCAGCCCGCCATGCTGAGCCTCCCTCCATGGGCTCTGTGCGGCCGAGCC 2369
|||||
Db 421 GGGACCTGCAGCCCGCCATGCTGAGCCTCCCTCCATGGGCTCTGTGCGGCCGAGCC 480
|||||
QY 2370 TCCCGAGCAGCAGCAGCCCTGCTCCACAGCGGCCAGTCCCATGCAACAGCAGAGGCT 2429
|||||
Db 481 TCCCGAGCAGCAGCAGCCCTGCTCCACAGCGGCCAGTCCCATGCAACAGCAGAGGCT 540
|||||
QY 2430 GAGAAGTGGGGGCGCAGCGCACCGGGAGCTGGCAGGAGCTAACCCCTGCAAGCCCTGGTGG 2489
|||||
Db 541 GAGAAGTGGGGGCGCAGCGCACCGGGAGCTGGCAGGAGCTAACCCCTGCAAGCCCTGGTGG 600
|||||
QY 2490 GAATCCACTGGGTGAAGCAGCTGGGCTCCTGAGTGTGGTGAAGTGGAGAACCTTTA 2549
|||||
Db 601 GAATCCACTGGGTGAAGCAGCTGGGCTCCTGAGTGTGGTGAAGTGGAGAACCTTTA 660
|||||
QY 2550 TGTCTAGCTCAGGGATGCTAAATACACCAATCAGCAGCCCTGTCTAGCTCAGGGTCTGT 2609
|||||
Db 661 TGTCTAGCTCAGGGATGCTAAATACACCAATCAGCAGCCCTGTCTAGCTCAGGGTCTGT 720
|||||
QY 2610 GAATGACCAATTCACACTCTGTATCTAGTACTGATGGGCGCTTGGAGAACCTTTAT 2669
|||||
Db 721 GAATGACCAATTCACACTCTGTATCTAGTACTGATGGGCGCTTGGAGAACCTTTAT 780
|||||
QY 2670 GTCTAGCTCAGGGATGCTAAATACACCAATCGGCACTGTATCTAGCTCAAGGTTGTA 2729
|||||
Db 781 GTCTAGCTCAGGGATGCTAAATACACCAATCGGCACTGTATCTAGCTCAAGGTTGTA 840
|||||
QY 2730 AACACCAATTCAGCAGCCCTGTGTCTAGCTCAGGGTATGTGAATGCACCAATGAGAGTC 2789
|||||
Db 841 AACACCAATTCAGCAGCCCTGTGTCTAGCTCAGGGTATGTGAATGCACCAATGAGAGTC 900
|||||
QY 2790 TGTATCTGCTACTTTCATGAGGCACTCGTGTGAAGAGACACCAACAGGCTTGTGTA 2849
|||||
Db 901 TGTATCTGCTACTTTCATGAGGCACTCGTGTGAAGAGACACCAACAGGCTTGTGTA 960
|||||
QY 2850 GCAATTAAGCTTCTATACCTGGGTGACAGTGGGCTGAGTCCGAAAAAGAGTCAAGCA 2909
|||||
Db 961 GCAATTAAGCTTCTATACCTGGGTGACAGTGGGCTGAGTCCGAAAAAGAGTCAAGCA 1020
|||||
QY 2910 GGGAGATTAAGGTTGGGGCGCTTTTATAGATTTGGGTAGGTAAAGAAAAATACAGTCA 2969
|||||
Db 1021 GGGAGATTAAGGTTGGGGCGCTTTTATAGATTTGGGTAGGTAAAGAAAAATACAGTCA 1080
|||||
QY 2970 AGGGGGTTTGTCTGCGGGGCGAGAGTGGGGGTGCGAAGTGTCTCAGTGGGGTGT 3029
|||||
Db 1081 AGGGGGTTTGTCTGCGGGGCGAGAGTGGGGGTGCGAAGTGTCTCAGTGGGGTGT 1140
|||||
QY 3030 TTTTGAAGCAGATGAGCCAGAGAAAAAGACTTTTCAACAAGTATGTCAATTAAGCA 3089
|||||
Db 1141 TTTTGAAGCAGATGAGCCAGAGAAAAAGACTTTTCAACAAGTATGTCAATTAAGCA 1200
|||||
QY 3090 AGGACCCGCGCATTTACCTCTTTTGTGTGAATGTCAAGTTAAGTTGGGCGAGGCT 3149
|||||
Db 1201 AGGACCCGCGCATTTACCTCTTTTGTGTGAATGTCAAGTTAAGTTGGGCGAGGCT 1260
|||||
QY 3150 AATATTCACCTTTTGTGTGTCTTCACTTACTTCAAGGCATCTGGGCGATATGTGCAAG 3209
|||||

Db 1261 AATATTCACCTTTTGTGTGTCTTCACTTACTTCAAGGCATCTGGGCGATATGTGCAAG 1320
|||||
QY 3210 TTTACAGGGATGCGATGCTTGGCTTGGGCTCAGAGGCTTGACAGCTACTTGTGTGGGCT 3269
|||||
Db 1321 TTTACAGGGATGCGATGCTTGGCTTGGGCTCAGAGGCTTGACAGCTACTTGTGTGGGCT 1380
|||||
QY 3270 CTTGGAAGTGTGTGTGTGACACTGTGTATCTAGTTAATCTAGTGGGACGTGGAAC 3329
|||||
Db 1381 CTTGGAAGTGTGTGTGTGACACTGTGTATCTAGTTAATCTAGTGGGACGTGGAAC 1440
|||||
QY 3330 CTTGTGTAGCTAGGGATTGTAAAGCAGCAATCAGGCGCTGTCAAAAAGACAGCAG 3389
|||||
Db 1441 CTTGTGTAGCTAGGGATTGTAAAGCAGCAATCAGGCGCTGTCAAAAAGACAGCAG 1500
|||||
QY 3390 TCGGCTCTACCAATCAGCAGATGTGGTGGGCGCAGATTAAGAAATAAAGAGGCTGC 3449
|||||
Db 1501 TCGGCTCTACCAATCAGCAGATGTGGTGGGCGCAGATTAAGAAATAAAGAGGCTGC 1560
|||||
QY 3450 CCGAGCCAGCAGTGGCAGACGGCGCAGGTCCCTATCCACAAATATGGCAGCTTGTCTTT 3509
|||||
Db 1561 CCGAGCCAGCAGTGGCAGACGGCGCAGGTCCCTATCCACAAATATGGCAGCTTGTCTTT 1620
|||||
QY 3510 TCGTGTGTGATTAATCTTGTCTACTGCTCGCTTTTGGGTCCACACTGCTTTATAGAG 3569
|||||
Db 1621 TCGTGTGTGATTAATCTTGTCTACTGCTCGCTTTTGGGTCCACACTGCTTTATAGAG 1680
|||||
QY 3570 TGTACACTCACACAGAGGTCTGACGCTTCACTCTGAAAGCCACTTAAGACAGAGGCC 3629
|||||
Db 1681 TGTACACTCACACAGAGGTCTGACGCTTCACTCTGAAAGCCACTTAAGACAGAGGCC 1740
|||||
QY 3630 ACCGGAGGAATTAACAACTCCGGCGCGGCTGCTTAAAGCTATTAACCTCAGCCGAA 3689
|||||
Db 1741 ACCGGAGGAATTAACAACTCCGGCGCGGCTGCTTAAAGCTATTAACCTCAGCCGAA 1800
|||||
QY 3690 GGTCTGACGCTTCACTCTCAGCAGCAGAGACACAGCAACCCAGAAAGAAAGAACTGC 3749
|||||
Db 1801 GGTCTGACGCTTCACTCTCAGCAGCAGAGACACAGCAACCCAGAAAGAAAGAACTGC 1860
|||||
QY 3750 GAACACATCTGAACATCAGAGAGAACAACTCCAGATGACACCACTTAAAGCTGTAA 3809
|||||
Db 1861 GAACACATCTGAACATCAGAGAGAACAACTCCAGATGACACCACTTAAAGCTGTAA 1920
|||||
QY 3810 CTCACCTGCGAGGGTCCGGCGCTTCTTGAAGTGAAGTGAAGACCAAGCACTCACAGTT 3869
|||||
Db 1921 CTCACCTGCGAGGGTCCGGCGCTTCTTGAAGTGAAGTGAAGACCAAGCACTCACAGTT 1980
|||||
QY 3870 TCGAGACAGAGCCAGAGATTGTAGATCAGCCGCGGCAACATGATGAATGCCCTCTG 3929
|||||
Db 1981 TCGAGACAGAGCCAGAGATTGTAGATCAGCCCTGGGCAACATGATGAATGCCCTCTG 2040
|||||
QY 3930 CAAAAAATTTTACAAAAATTTGGGGAAGCATGTGTCCCTGCTGTGTGCCAG 3989
|||||
Db 2041 CAAAAAATTTTACAAAAATTTGGGGAAGCATGTGTCCCTGCTGTGTGCCAG 2100
|||||
QY 3990 CTACCGGGAGGCTTAAAGTGGAGGATCGCTTGATGAGCTGGAGGTGAAGCTCAGTGA 4049
|||||
Db 2101 CTACCGGGAGGCTTAAAGTGGAGGATCGCTTGATGAGCTGGAGGTGAAGCTCAGTGA 2160
|||||
QY 4050 CTGTGATTTACACAGCCCTGAGGCTGGGGACAGCATGAGAACCTTTCCTCCCTCG 4109
|||||
Db 2161 CTGTGATTTACACAGCCCTGAGGCTGGGGACAGCATGAGAACCTTTCCTCCCTCG 2220
|||||
QY 4110 CAAAAAATTTGACAAAAATGTATAAGAGGTGCTGTATATGCTTAGGCGCAGTGCCTAT 4169
|||||
Db 2221 CAAAAAATTTGACAAAAATGTATAAGAGGTGCTGTATATGCTTAGGCGCAGTGCCTAT 2280
|||||
QY 4170 GCTGTATATCCAGCAGCTTTGGGAAGCCAGAGCGGGGCTCACTTAAGTCAAGAGTGT 4229
|||||
Db 2281 GCTGTATATCCAGCAGCTTTGGGAAGCCAGAGCGGGGCTCACTTAAGTCAAGAGTGT 2340
|||||
QY 4230 GAGACAGGCTGGCCAACTGAGAAAGCCCATCTCTTATAAATTAACAAATTTAGCCGG 4289
|||||

Db	2341	GAGACCA	GCC	TGGCCAA	CA	TG	GAGAAAG	CCCA	TC	TTCTTTAA	AAATTC	CAAAAT	TAGCCGG	2400
QY	4290	CTGTGGG	GGC	AGTGGT	GG	AGC	ATCC	CTGTAA	TCC	CAGCTAC	TC	GAGAGG	CTGAGGCAG	4349
Db	2401	CTGTGGG	GGC	AGTGGT	GG	AGC	ATCC	CTGTAA	TCC	CAGCTAC	TC	GAGAGG	CTGAGGCAG	2460
QY	4350	GAATC	ACTTGA	AA	CCCA	GAGAGG	CGGGG	GGTGG	AG	GAGCC	CA	GCATG	CCACTCC	4409
Db	2461	GAATC	ACTTGA	AA	CCCA	GAGAGG	CGGGG	GGTGG	AG	GAGCC	CA	GCATG	CCACTCC	2520
QY	4410	ACCCAC	TC	CCAGC	CC	TGGG	CA	CA	CAG	CCAA	CC	TGTCTT	AAAA	4469
Db	2521	ACCCAC	TC	CCAGC	CC	TGGG	CA	CA	CAG	CCAA	CC	TGTCTT	AAAA	2580
QY	4470	CCTGAC	AT	ATA	TAA	AGG	GT	CG	AA	TG	TC	CCAG	CA	4529
Db	2581	CCTGAC	AT	ATA	TAA	AGG	GT	CG	AA	TG	TC	CCAG	CA	2640
QY	4530	GCTCC	CTC	CT	CCAT	TGGT	CT	CGT	TAA	AA	AC	CC	CA	4589
Db	2641	GCTCC	CTC	CT	CCAT	TGGT	CT	CGT	TAA	AA	AC	CC	CA	2700
QY	4590	GCC	TA	TAA	TCC	CAG	CA	CTT	TGG	GAG	CGC	CA	GAGG	4649
Db	2701	GCC	TA	TAA	TCC	CAG	CA	CTT	TGG	GAG	CGC	CA	GAGG	2760
QY	4650	GAGAC	CA	G	CGC	TG	AC	CA	CA	AT	TGT	GA	AAAT	4709
Db	2761	GAGAC	CA	G	CGC	TG	AC	CA	CA	AT	TGT	GA	AAAT	2820
QY	4710	TGAGC	AT	GT	GGT	G	T	G	T	G	T	G	T	4773
Db	2821	TGAGC	AT	GT	GGT	G	T	G	T	G	T	G	T	2834

```

RESULT 6
US-09-305-384-7
; Sequence 7, Application US/09305384
; Patent No. 6242218
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
; FILE REFERENCE: 07236/017001
; CURRENT APPLICATION NUMBER: US/09/305,384
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,649
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-305-384-7

```

Query Match	20.1%;	Score 1252;	DB 3;	Length 1252;
Best Local Similarity	100.0%;	Pred. No. 9.2e-292;		

QY	4728	CCTTAAATCCACACTACTTTGGGAGGCTGAGGCGAGAAAAATACATAGAACCAAGGAGGCGG	4787
Db	1	CCTTAATCCACACTACTTTGGGAGGCTGAGGCGAGAAAAATACATAGAACCAAGGAGGCGG	60
QY	4788	AGGTTGTAAGTAGCCGAGATGTGCCATTGTCACATCCAGCTGAGCAATGAGCGAAATCC	4847
Db	61	AGTTGTAGTAGGCGGAGATGTGCCATTGTCACATCCAGCTGAGCAATGAGCGAAATCC	120
QY	4848	ATCTCAAAAAACAACAACAAAAACCACTCTTACTCCAGGGAGCTGGGTACAGAGCT	4907
Db	121	ATCTCAAAAAACAACAACAAAAACCACTCTTACTCCAGGGAGCTGGGTACAGAGCT	180

[illegible]

RESULT 7
US-09-422-576D-1
; Sequence 1, Application US/09422576D
; Patent No. 6395549
; GENERAL INFORMATION:
; APPLICANT: Tuan, Dorothy
; APPLICANT: Long, Qiaoming
; APPLICANT: Bengra, Chikh
; TITLE OF INVENTION: Long Terminal Repeat, Enhancer, and Insulator Sequences for Use
; FILE REFERENCE: M0351-205010
; CURRENT APPLICATION NUMBER: US/09/422,576D
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/105,256
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-422-576D-1

Query Match 10.7% Score 669.2; DB 4; Length 1831;
Best Local Similarity 68.0%; Pred. No. 2,2e-151;
Matches 1396; Conservative 0; Mismatches 248; Indels 408; Gaps 17;

QY 1873 TGAGAGGTGACAGCATGCGCGCAGTCTCAGACGCCCTGTCGCTCTGCGGCGCTCTCT 1932
DB 5 TGAGAGGTGACAGCATGCGCGCAGTCTCAGACGCCCTGTCGCTCTGCGGCGCTCTCT 64
QY 1933 GCTGGGGCTCCACTTGGTGGCACTGTAGAGAGCCCTTACAGCCCAACCCCTGCTAGTGGG 1992
DB 65 GCTGGGGCTCCACTTGGTGGCACTGTAGAGAGCCCTTACAGCCCAACCCCTGCTAGTGGG 124
QY 1993 AGCCCTTTCTGGGGTGGCCAGAGCGCGCTCCCTCAGTCTGTAGAGAGGTGG 2052
DB 125 AGCCCTTTCTGGGGTGGCCAGAGCGCGCTCCCTCAGTCTGTAGAGAGGTGG 184
QY 2053 AGGAGAGGCTCAGCAGAGAACCGGGGCTGCGCACGCGGCTTGGGGCAGCTGAGTTC 2112
DB 185 AGGAGAGGCTCAGCAGAGAACCGGGGCTGCGCACGCGGCTTGGGGCAGCTGAGTTC 235
QY 2113 CGGGGGGCGTGGGGCTTGGGGGCGCGCAGTCCGAGAGAGCGGGCAGCCCTGAGAGCC 2172
DB 236 CGGGGGGCGTGGGGCTTGGGGGCGCGCAGTCCGAGAGAGCGGGCAGCCCTGAGAGCC 294
QY 2173 CGGGGCAATGAGAGGCTTAGACACCGGGGCTGCGGAGGGGTGTAAGTGGTGCC 2232
DB 295 GCGGGCAATGAGAGGCTTAGACACCGGGGCTGCGGAGGGGTGTAAGTGGTGCC 335
QY 2233 AGCAGTGCACGCCGCCGCGCTGTGCTCGATTCGATTCGATTCGATTCGATTCGATTC 2292
DB 336 -----TGCTCAATTCCTGCGCCGCGGCTTAGTGGCTT 367
QY 2293 CCCGGGGGCGAGGGCTGGGAGCTTGCAGACCGCGCAATGCTGAGCTTC-----CTTCAGATG 2349
DB 368 CTTGGGGGCGAGGGCTGGGAGCTTGCAGACCGCGCAATGCTGAGCTTCCTCCACCTTCATGG 427
QY 2350 GCTCTGTGCGCGCGAGGCTTCCCGAGAGCAGCAGCCCTGCTGCTCAGACGCGCCATG 2409
DB 428 GCTCTGTGCGCGCGAGGCTTCCCGAGAGCAGCAGCCCTGCTGCTCAGACGCGCCATG 487
QY 2410 CCATTCAGACGCAAGGGCTGAGAAATGCGGGGCGACGCGCAGCTGGAGCTGGCAGGCACT 2469
DB 488 CCATTCAGACGCAAGGGCTGAGAAATGCGGGGCGC -CAGCAAGGGGACTGGCAGGCACT 546
QY 2470 ACCCTTCAGAGCCCTGTGCGGAATTCAGTGGGTGAAGCAGCTGGGCTCTGAGTGTGT 2529
DB 547 CCCCCCTGAGGCCAGGTGTGCGGATTCAGTGGGTGAAGCAGGCTGAGG -TCTGAGATTTGCT 605
QY 2530 GGAGACTTGGAGACCTTATGTAGTCAAGGGGATGCTAAT----- 2572

DB 606 GGGGATGCGAAGAACCTTATGTCTAGATTAAGGATTTGTAATACCAATTTGGCACTCT 665
QY 2573 -----ACACAATACACACCTGTGTCTAGTCTAGGCTGCG 2608
DB 666 GTATCTAGCTCAAGTTTGTAAACACACCAATACACACCTGTGTCTAGGCTGCG 725
QY 2609 TGAATGCAACCAATCAGACTGTATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 2668
DB 726 TGAATGCAACCAATCAGACTGTATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 783
QY 2669 TGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCTGTATCTAGTCTAGTCTAGTCTAG 2728
DB 784 TGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCTGTATCTAGTCTAGTCTAGTCTAG 843
QY 2729 AAACACACCAATCAGACCTGTGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCT 2788
DB 844 AAACACACCAATCAGACCTGTGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCT 903
QY 2789 CTGTATCTAGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCTGTGTGTG 2848
DB 904 CTGTATCTAGTCTAGTCTAGTCTAGGATTTGTAATTAATACCAATTCGAGCTGTGTGTG 920
QY 2849 AGCAATTAAGCTTCTATACCTGCGGTGAGAGTGGGCTGATCCGAAAGAGATCAGGA 2908
DB 921 ----- 920
QY 2909 AGGAGATTAAGGCTGCGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2968
DB 921 -----GATGAGAGCTGAGAACTTTATGTCTAGTCTAGGATTTGTAATTAATTAAT 969
QY 2969 AAGGGGCTTTGTTCTGCGCGGCGAGAGTGGGGGTGCGCAAGTGTCTAGTGGGGGTGCG 3028
DB 970 ----- 969
QY 3029 TTTTGAACAGATGAGCAGCAAGAAAGACTTTCACAGGTAATGTATCATCAATTAAGC 3088
DB 970 -----CCACTGCGCAGTGTGTATCTAGTCTAGGATTTGTAATTAATTAATTAAT 1011
QY 3089 AAGGACCGCATTTATACCTCTTTTGTGTAATTAATTAATTAATTAATTAATTAATTAAT 3148
DB 1012 CAATGAGCAGCTGTGTCTAGTCTAGGATTTGTAATTAATTAATTAATTAATTAATTAAT 1059
QY 3149 CATATTCATCTTTTGTGATTTCTAGTCTAGTCTAGGATTTGTAATTAATTAATTAAT 3208
DB 1060 ----- 1059
QY 3209 GTTACAGGAGATGCGATGCTTGGCTGAGAGCTTGCAGAGCTTGCAGAGCTTGCAGAG 3268
DB 1060 -----CACTCTGTATCTAGTCTAGTCTAGGATTTGTAATTAATTAATTAATTAAT 1086
QY 3269 CTTTGGAGATTTGTGTCAGACTGTGTATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 3328
DB 1087 ACTTGGAGATTTGTGTCAGACTGTGTATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 1146
QY 3329 CTTTGGATTTGTGTCAGAGGATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3388
DB 1147 CTTTGGATTTGTGTCAGAGGATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1206
QY 3389 CTTGCG--CTTACCAATCAGAGATTTGGGTTGGGGCAGATTAAGAGATTAAGAGCAGC 3446
DB 1207 CTTGCGCTTCTACCAATCAGAGATTTGGGTTGGGGCAGATTAAGAGATTAAGAGCAGC 1266
QY 3447 TGCAGGACGAGAGTGGCAACGCGCAGAGTCCCTATCCAAATATAGGAGCTTTGTTTC 3506
DB 1267 TGCAGGACGAGAGTGGCAACGCGCAGAGTCCCTATCCAAATATAGGAGCTTTGTTTC 1326
QY 3507 TTTTGTGTTTGGCAATTAATTTGCTACTGCTGCTTTTGGTCCACACTGCTTTTATG 3566
DB 1327 TTTTGTGTTTGGCAATTAATTTGCTACTGCTGCTTTTGGTCCACACTGCTTTTATG 1386
QY 3567 AGCTGTACACTCAGCAAGAGTGTGAGCTTCACTCTCAAGCAGCTTAAGACAGAG 3626
DB 1387 AGCTGTACACTCAGCAAGAGTGTGAGCTTCACTCTCAAGCAGCTTAAGACAGAG 1446

```

QY 3627 CCCACGGAGATGAACTCCGGCGCGCTGCTTAAAGGCTATACACTACCGC 3686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 CCCACC -GAGAGAAAGAACTCCAGAGGCGCG -CTTAAGAGCTGGAACTTACTCT 1504
QY 3687 GAAGGTCTGCAGCTTCTACTCTCAGCAGCAGACGAGAACCCACGAAGAAAGAAC 3746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1505 GAAGGTCTGCAGCTTCTACTCTCAGCAGCAGACGAGAACCCACGAAGAAAGAA-- 1562
QY 3747 TCGCAACACTGTGAATCAGAGAGAACTCCAGATGACCACTTAAAGCTGTGA 3806
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1563 CTCGAACACTCAACATCAGAACTCAGAACTCCACACAGCAGCTTTAAGAACTGTA 1622
QY 3807 ACACACTGCGAGGCTCCGGGCTTCTCTTGAAGTACAGAGACCAAGCACTACCA 3866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1623 ACACACTGCGAGGCTCCGGGCTTCTCTTGAAGTACAGAGAACTCAAGAACCA 1682
QY 3867 GTTTCGGACACA 3878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1683 ATTCCGGACACA 1694

```

RESULT 8

```

US-09-814-951A-3
; Sequence 3, Application US/09814951A
; Patent No. 6387661
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al
; TITLE OF INVENTION: ISOLATED HUMAN AMINOACYLASE, NUCLEIC
; FILE REFERENCE: C1001179
; CURRENT APPLICATION NUMBER: US/09/814,951A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9704
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-951A-3

```

```

Query Match          9.0%; Score 563.8; DB 4; Length 9704;
Best Local Similarity 86.9%; Pred. No. 1.1e-125;
Matches 686; Conservative 0; Mismatches 77; Indels 26; Gaps 5;

```

```

QY 1873 TGAGAGGTGACAGCATGCCGCGAGTCTTCACAGCCCTTGTTGGCTTGGCGCTCTCT 1932
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8883 TGAGAGGTGACAGCATGCCGCGAGTCTTCACAGCCCTTGTTGGCTTGGCGCTCTCT 8942
QY 1933 GCCTGGGCTCCCACTTGGTGGCACTTGAAGAGGAGCCCTTCAGGCCACCGCTGACGTGG 1992
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8943 GCCTGGGCTCCCACTTGGTGGCACTTGAAGAGGAGCCCTTCAGGCCACCGCTGACGTGG 9002
QY 1993 AGCCCTTTTGGGGCTGGCCAAAGGCGAGAGCGGCTCCCTCAGCTTGCAGGAGGAGTGG 2052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9003 AACCCCTTTGGGGCTGGCCAAAGGCGAGAGCGGCTCCCTCAGCTTGCAGGAGGAGTGG 9062
QY 2053 AGGAGAGGCTCAACAGAGAACCGGGGCTGGCGACAGGGGCTTGGGGGCGCAGCTGAGTTTC 2112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9063 AGGAGAGGCTCAACAGAGAACCGGGGCTGGCGACAGGGGCTTGGGGGCGCAGCTGAGTTTC 2122
QY 2113 CGGGTGGGCGTGGGCTTGGGGGCGG-----CGCACTCGAGAGAGCGGGCGAGCCCTGAC 2166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9123 CGGGTGGGCGTGGGCTTGGGGGCGGCGCCGCACTCGAGACTCGAGAGAGCGGGCGCCCTGAC 2182
QY 2167 -----CAGGCCCGGGGCAATGAGAGCTTAGACCCCGGGCGAGCGGCTGGAGAGGT 2218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9183 CGTCCCGCGCTGCCCGGGGCAATGAGGGGCTTAGACCCGGGCGAGTGGCTCGGAGGGT 2242
QY 2219 GTACTGGGTGCCCGGAGAGTGGCAGCCCGCGGGGCTGTGCTCGCTTCTCTACTGG 2278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9243 GTACTGGGTGCCCGGAGAGTGGCAGCCCGGCGCTG---CTCTCGATTTCTCACCGG 9298

```

```

QY 2279 GCCTTACAGCCTTCCCGGGGAGAGGCTCGGAGCTTCAGCCCGCCATGCT----- 2332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9299 GTCTTAGCTGCTTCCCGGGGCTCAGGGTTTGGAGCTTCAGCCACCACTGCTTAGAGCC 9358
QY 2333 -GAGCTCCCTTCATGGGCTCTGTGGCGCCGAGCCTCCCGAGACGACACACCCCT 2391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9359 CTCGCCACCCCTTCAGCTGCTCCGTTGGGGCCCAAGCCTCCCCCAATGAGCGCGCCGCC 9418
QY 2392 GCTCCAGAGCGCCGAGCTCCATTCAGACACGAGAGGCTTGAAGATGCGGGGCGCAGGCA 2451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9419 GCTCCAGCGCACAGTCTCCATTCACACCCCAAGGAGCTGAGAGTGCAGGCTCAGG -AG 9477
QY 2452 CGGGAGCTGGGAGAGCTACCCCTGAGCGCCGTGGCGGAATCCATGGGTGAAGCCAGC 2511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9478 CAGAGCTGGAGGAGCTCAGCTCAGCGCCCGGTGGGATTCATGGGTGAAGCCAGC 9537
QY 2512 TGGGCTCCAGTCTGGTGGAGAGCTTGAAGAACTTATGTAGCTCAGGAGTGTAA 2571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9538 TGGGCTCCAGTCTGGTGGAGAGCTGAGAACTTATGTAGCTCAGGAGTGTAA 9597
QY 2572 TACACCAATCAGACCCCTGTCTAGCTCAGGCTGTGAATGACCAATCCACTCTG 2631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9598 TACACCAATCAGACCCCTGTCTAGCTCAGGCTGTGAATGACCAATCCACTCTG 9657
QY 2632 TATCTAGCT 2640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9658 TCTCTAGCT 9666

```

RESULT 9

```

US-09-422-576D-5
; Sequence 5, Application US/09422576D
; Patent No. 639549
; GENERAL INFORMATION:
; APPLICANT: Tuan, Dorothy
; APPLICANT: Long, Qiaoming
; APPLICANT: Bengra, Chikh
; TITLE OF INVENTION: Long Terminal Repeat, Enhancer, and Insulator Sequences for US
; FILE REFERENCE: M0351-205010
; CURRENT APPLICATION NUMBER: US/09/422,576D
; CURRENT FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/105,256
; PRIOR FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1091
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus Sequence
; NAME/KEY: misc_feature
; LOCATION: (437)..(437)
; OTHER INFORMATION: n = any nucleotide
US-09-422-576D-5

```

```

Query Match          7.4%; Score 462.2; DB 4; Length 1091;
Best Local Similarity 88.8%; Pred. No. 1.2e-101;
Matches 522; Conservative 0; Mismatches 64; Indels 2; Gaps 2;

```

```

QY 3293 CTCGTATCTATCTTATCTAGTGGGAGCTGAGAACTTGTGCTACTCAGGAGTTG 3352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 410 CTCAGTATCTATCTTATCTAGTGGGAGCTGAGAACTTGTGCTACTCAGGAGTTG 469
QY 3353 TAAAGCAACCAATCAGCGCCCTGTCAAAAACAGACCACTCGGCTC-TACCAATCAGCAGA 3411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 470 TAAATGCAACCAATCAGTGGCTGTCAAAAACAGACCACTCGGCTCTTACCAATCAGCAGA 529
QY 3412 TGTGGGTGGGGCGACATAAAGATPAAAGCAGGCTGCCGAGCCAGCAGTGGCAACGG 3471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 530 TGTGGGTGGGGCGACATAAAGATPAAAGCAGGCTGCCGAGCCAGCAGTGGCAACCGG 589

```


Db	327	CTTTGTTCTTTGCGCTTTGGAAATTAATCTTGGCTGCTGCTACAGTATTGGGCTTACACTG	386
OY	3559	CTTTTATGAGCTGTATACACTATCACCAGGAAGTCTGCAGCTTCACTCCTGAAGCACTAAG	3618
Db	387	CTTTAGAGCTAATAAGCTCACC -CGAAGGTCCTGACGCTTCACTCTTGAAGCCAGCGAG	445
OY	3619	ACCAGAGCCCAACCGGAGAGAAATTAACAACCTCCGGCCGGCTTGCTTAAAGCTATATA	3678
Db	446	ACCAGAAACCCACGTGGGAGGAACGAAACAACCTCCAGAGCAGCCGCTTAAAGAGCTGAACG	505
OY	3679	CTCACCGGAAAGTCTGCAGCTTCACTCCTTACGACGAGAGAACGAAACCAACCAACG	3738
Db	506	TTCACTGTGAAGGCTCTGACGCTTCACTCTTAAGCCAGCAGAGAACCAAGCAACCACTATGAGAG	565
OY	3739	GAAGAACAATCGGAACACATCTGAACATCTAGAAGGAACAACCACTCCAGATGACCAAC -TTA	3797
Db	566	GAAGAACAATCGGAACACATCTGAACATCTAGAAGGAACAACCACTCCAGATGACCAACGCTT	625
OY	3798	AGAGCTGTAACTCACTGCGAGGGT -CCGCGGCTTCTCTTG -AAGTAGTAGACCA	3855
Db	626	AGAACTGTAACTCACTGCGAGGGTCCGCGGCTTCACTTCTTGAAGTAGTAGTAACCA	685
OY	3856	AGCACTCAACGAGTT	3869
Db	686	AGAACTTACCAATT	699

```

RESULT 12
US-09-422-576D-6
; Sequence 6, Application US/09422576D
; Patent No. 6395549
; GENERAL INFORMATION:
; APPLICANT: Tuan, Dorothy
; APPLICANT: Bengta, Chikh
; APPLICANT: Bengta, Qiaomng
; TITLE OF INVENTION: Long Terminal Repeat, Enhancer, and Initiator Sequences for Use i
; FILE REFERENCE: M0351-205010
; CURRENT APPLICATION NUMBER: US/09/422,576D
; CURRENT FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/105,256
; PRIOR FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-422-576D-6

```

Query Match	6.88	Score 426.4	DB 4	Length 1043
Best Local Similarity	88.5%	Pred. No. 4.9e-93		
Matches 485	Conservative 0	Mismatches 61	Indels 2	Gaps 2
QY	3333	TGTCGTACGCTCAGGAGTGTGTAACGGCACCAATAGAGGCCCTGTCCAAACAGACACTCG	3392	
Db	402	TGTCGTACGCTCAGGAGTGTGTAACGGCACCAATAGAGGCCCTGTCCAAACAGACACTCG	461	
QY	3393	GCTCTACCAATCAGCAGAGATGTGGGTGGGGCCAGATTAAGAAATAAAAGCAGGCTGCCG	3452	
Db	462	GCTCTACCAATCAGCAGAGATGTGGGTGGGGCCAGATTAAGAAATAAAAGCAGGCTGCCG	521	
QY	3453	AGCCAGCAGTGGCAACGCGCACAGGTCCCTAATCCACATATATGCGAGCTTGTCTTTTGC	3512	
Db	522	AGCCAGCAGTGGCAACGCGCGGTCCCTCCACACTGTGGAAGCTTGTCTTTTGC	581	
QY	3513	TGTTTGGGATTAATCTTGTCTACTGCTGCTTTTGGGTCCACACTGCTTTTATAGCTGT	3572	
Db	582	TCTTTGCAATTAATCTTGTCTGCTCTCACTGTTTGGGTCTCACTGCTTTATAGCTGT	641	
QY	3573	AACATCAACACGAAGGTGTGCACCTTCACCTCCGAAGCCACTTAAGACACAGACCAC	3632	
Db	642	AACGCTCAACGCGGAAGGTGTGCACCTTCACCTTTGAAGCCAGCGACACGAACAC	701	

QY	3633	GGGAGGAATGAACAACTCCGGCCGGCGCTTTAAGAGCTATAACACTCAC- GGCGAAG	3691
Db	702	GGGAGGAACGACAACTCCAGAGCGCGCTTTAAGAGCTTGAAGCTTCAAGG	761
QY	3592	TCTACAGCTTACTCTCTCAGGCAGCAGACACGACCAACCCACAGAAAGAAACTCGCA	3751
Db	762	TCTCAGAGTTCACTCTCTGAGCCAGCAGACCAACCAATCAGAAAGAAACTCGCA	821
QY	3752	ACACATCTGAAACATCAGAAAGAAACCAACTCCAGATGCAACCAAC--TTAAGAGCTTAAAC	3810
Db	822	ACACATCCAAACATCAGAAAGCAACCAACTCCACACAGCAGACCTTTAAGAACTTAACAC	881
QY	3811	TCACCTGGAGGGTCCGGGCTTCTCTCTTGAAGTCAGTGAAGACCAACACTCAGACAGTTT	3870
Db	882	TCACCAAGAGGGTCCCGGCTTCACTTGTGAAGTCAGTGAAGAAACCAAGAACCCACCAATTC	941
QY	3871	CGGACACA	3878
Db	942	CGGACACA	949

```

RESULT 13
; US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078.294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0.
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
; US-09-078-294-4

```

Query Match	6.63;	Score 413.8;	DB 3;	Length 80246;
Best Local Similarity	66.08;	Pred. No. 3,7e-89;		
Matches 806;	Conservative	0;	Mismatches 357;	Indels 58; Gaps 12;
QY 3259	TCTGTGGGGCTTGGAGATGTTTGTGCGACACTGTATCTATGTTATCTAGTGGGG	3318		
Db 59197	TCAGGTGGGACTTGGAGAACTTTTGTGCTC-----TACTTAAGATTGTAAATGC	59247		
QY 3319	ACGTGGAGAACTTTGTGTCTAGCTCAGGGATTGAAACGACCAATCAGCGCCCTGTCA	3378		
Db 59248	ACCAATCAGCACTCTGTCTAGCTTAAAGATTGAAATGACCAATCAGCACTGCTGTAA	59307		
QY 3379	AAACAGACCACTGGGCTTACAATGAGAGATGGGTGGGGCAGATTAAGAAATA	3438		
Db 59308	AA-----TGGHCAATAGAGATGGGGGGGGTCAAAATTAAGGAGTAA	59353		
QY 3439	AAGCAGCGTCGCCGACGACAGAGTGGCAACGGCAGACAGTCCCTATCCACAATATGCGAG	3498		
Db 59354	AAACTGGCCACCCGACGACAGAGTGGCAACCCACTGGGGTCCCTTCCACACTGCGAAG	59413		
QY 3499	CTTTGTCTTTTGTCTTTTGGCATTAATTTTCTCTACTGTGCGCTTTTGGGTCCACACTG	3558		
Db 59414	CTTTGTCTTTTGTCTCTTACACATAAATCTTCTCTCTGTCTCAATCTTTGTGTGCCACTTA	59473		
QY 3559	CTTTTATGAGCTTAACTACTCAACAGAGTCTGCAGCTTCTACTCTGAAACCACTAAG	3618		
Db 59474	CTTTTATGAGCTTAACTACTCACTGAGAGGCTGTGGGCTTATTTCTGAAATCA-ACAG	59532		
QY 3619	ACCAGAGCCACCCGGGAGAAATGAACAATCCGGCGCGCTGCC-TTATAGAGTATAAC	3677		
Db 59533	ACCAACCAACCACTGGAGAGAACCAAGAACTCCCATGTGCTGCTTTAAAGAGCTATAAC	59592		

QY 3678 ACTCAGCGAAGGCTGTCAGACTTCTACTCT-CAGCCAGCGAGACCAAGACCCACAGAGA 3736
 Db 59593 ACTCAGCGAAGGCTGTCAGACTTCTACTCTGAGTCAAGTGAAGACCAACCAACCCACAGAGA 59552
 QY 3737 AGGAAGAACTGGCAACATCTGAACATCGAAGGAAGAAACCACTCCAGATGCACCA-CCT 3795
 Db 59653 AGGAAGAACTGGCAACATCTGAACATCTGAAGGAAGAAACCACTCCAGATGCACCT 59712
 QY 3796 TAAAGCTGTAACTACTGTCGAGGGTCCGCGCTTCTCTTGAAGTCAAGTGAAGCA 3855
 Db 59713 TCAGAGCTGTAACTACTGTCGAGGGTCCGCGCTTCTCTTGAAGTCAAGTGAAGCA 59772
 QY 3856 AGCACTACCACTGTCGAGCAACAAGCCAGAGATTGAGATCAGCTGGGCAACATGATG 3915
 Db 59773 AGAAGCCAGCGAAGGAACAATTCAGACACAGTAGAATCTGTATTTTGTGATCTGTG 59832
 QY 3916 AATGACCTCTCTCCAAAAAATAATCAAAAAATTTGGCCGAGCATGGTCTCCGT 3975
 Db 59833 GCTTCCAGGCTTACTCCAGTCTATTGAAGTCTTCCATTTCAGCCCTTAAGGAACAGAGATG 59892
 QY 3976 GCTGTGTGTCCAGCTACGCGGGAGGCTAAAGTGGAGATCGCTTGAGCCT-GGAGGT 4034
 Db 59893 GTTGTGGAGAGACATGTGGAAATTTGTTATGGACAGGCTTGAGATGCACATAGGGCAT 59952
 QY 4035 GAAGACTGCACTGAGCTGTGATTGTACCAAGCCCTTACGCTGGGGACAGACT----- 4089
 Db 59953 TCTGTCAAACTCTGAGCTGGAGAGAGGCGCAGAAATTAATCTAAGGAAGACAGTCTTTTG 60012
 QY 4090 ----GAGACCTGTTCCTCCCGCAAAAAAATGACAAAAATGTAATTAAGAGTGGCTG 4145
 Db 60013 TAGACAGTAGTAGTCTTGTGCTTGTGAGACATGTAGATTATCAACCAATTATTAAGAAAA 60072
 QY 4146 ATATGCTAGGCGAGTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCGAGGGG 4205
 Db 60073 ATATAGCAGGTGGAGTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCGAGGGTG 60132
 QY 4206 CGGGTACCTTAAGGTCAGAGATGTGAGACCAAGCCTTGGCCACATAGGAAGCCCATCTC 4255
 Db 60133 TGGATCAC--AAGGTCAAGGCTTGTGAGACCAAGCCTTGGCCACATAGGAAGCCCATCTC 60190
 QY 4266 TTTCAAAAAATTAAGGTCAGGCTGTGGGAGCATGTGAGACATCTGTATCCCA 4325
 Db 60191 TACTAAAAATTAAGGTCAGGCTGTGGGAGCATGTGAGACATCTGTATCCCA 60242
 QY 4326 GCTTCTAGAGGCTGAGGCAAGAAATCACTTAACCCAGAGAGCGGGCTTGCAGTGA 4385
 Db 60243 G-TACTAGAGGCTGAGGCAAGGGAATCTCTGAACCTTGGAGGCAAGAGGTTGCAGTGA 60301
 QY 4386 GCCGAGATCGTGGCATTTGCACTCCAGCCACTCCAGGCGCAACAAAGAGCCAACTGTG 4445
 Db 60302 GCCAAGATCACACACAGCACTCCATCC-----TGGGTGACAGAGGAGACTCTG 60351
 QY 4446 TCTTAAAAAATAAAAAA 4466
 Db 60352 TCTCAAAAAAATAAAAAA 60372

LENGTH: 80595
 ; TYPE: DNA
 ; ORGANISM: Nucleotide sequence of HC-contig
 us-09-078-294-3
 Query Match 6.6%; Score 412.2; DB 3; Length 80595;
 Best Local Similarity 65.9%; Pred. No. 9e-89;
 Matches 805; Conservative 0; Mismatches 358; Indels 58; Gaps 12;
 3259 TCTGTGGGGGCTGGGAATGTGTGTGTCAGACATCTGTATAGTTAATCTAGTGGG 3318
 Db 59430 TCAGGTGGGACTTGGAGAACTTTGTGTCT-----TAGCTAAGGATTGTAATGC 59480
 QY 3319 ACGTGGAACCTTTGTGTGTAGCTCAGAGGATTGTAACGCAACATCAGGCCCTGTCA 3378
 Db 59481 ACCAATCAGACCTGTGTGTAGCTAAGGATTGTAATGACCAATCAGACATCTGTAA 59540
 QY 3379 AAACAGACACTCGGCTCTACCAATCAGCAGAGATGTGGTGGGCCAGTAAAGAAATTA 3438
 Db 59541 AA-----TGGACCAATCAGCAGAGATGTGGGCGGTCAAAATTAAGGAGTTAA 59586
 QY 3439 AAGCAGCTGCCGAGCCAGCAGAGTGGCAACGCGCACAGTCCATATCCAAATATGGCAG 3498
 Db 59587 AAATGGCCACCCGAGCCAGAGTGGCAACCCACTCGGTCCTTCCACACTGTGAAG 59646
 QY 3499 CTTTGTCTTTTGTCTGTGTGCGATTAATCTTGTCTAGTCTGCTGCTTTTGGTCCACTG 3558
 Db 59647 CTTTGTCTTTTGTCTGTGTGCGATTAATTAATCTTGTCTGCTGCTATCTTTGTGTCCACTA 59706
 QY 3559 CTTTATAGAGTGTAAACATCAGCAGAGTGTGACGCTCATCTCGTAAGCCACTAAG 3618
 Db 59707 CTTTATAGAGTGTAAACATCAGCAGAGTGTGACGAGGCTGTGTGCTTCAATCTTGAGTGA -ACAG 59765
 QY 3619 ACCAGAGCCAGCCGAGGAGATGAACACTCCGCGCGCTGCC-TTAAAGCTATATAC 3677
 Db 59766 ACCAGAGCCAGCTGGAAGGAACAAAGAACTCCGAGATGTGCTTTAAGAGCTGTAAAC 59825
 QY 3678 ACTCAGCGAAGGCTGTCAGACTTCTACTCT-CAGCCAGCGAGACCAAGACCCACAGAGA 3736
 Db 59826 ACTCAGCGAAGGCTGTCAGACTTCTACTCTCGAAGTGAAGTGAAGCAACCAACCCACAGAGA 59885
 QY 3737 AGGAAGAACTGGCAACATCTGAACATCTGAAGGAAGAAACCACTCCAGATGCACCA-CCT 3795
 Db 59886 AGGAAGAACTGGCAACATCTGAACATCTGAAGGAAGAAACCACTCCAGATGCACCACTCTT 59945
 QY 3796 TAAAGCTGTAACTACTGTCGAGGGTCCGCGCTTCTCTTGAAGTCAAGTGAAGCA 3855
 Db 59946 TCAGAGCTGTAACTACTGTCGAGGGTCTGTGCTTCAATCTTGAAGTCAAGAAAGCA 60005
 QY 3856 AGCACTACCACTGTCGAGCAACAAGCCAGAGATTGAGATCAGCTGGGCAACATGATG 3915
 Db 60006 AGAAGCCAGCGAAGGAACAATTCAGACACAGTAGAATCTGTATTTTGTGATCTGTG 60065
 QY 3916 AATGACCTCTCTCCAAAAAATAATTAATTAATTTGGCCGAGCATGGTCTCCGT 3975
 Db 60066 GCTTCCAGGCTTACTCCAGTATTTGAAGTCTTCCATTTCAGAGCTTTAAGGAACAGAGATG 60125
 QY 3976 GCTGTGTGTCCAGCTACGCGGGAGGCTAAAGTGGAGATCGCTTGAAGCT-GGAGGT 4034
 Db 60126 GTTGTGGAGAGACATGTGGAAATTTGTTATGGACAGGCTTGAAGTGAACATAGGGCAT 60185
 QY 4035 GAAGACTGCACTGAGCTGTGATTGTACCAAGCCCTTACGCTGGGGACAGACT----- 4089
 Db 60186 TCTGTATCAAACTCTGAGCTGGAGAGGCGCAGGAATTAATCTAAGGAAGCACTTTTGTG 60245
 QY 4090 ----GAGACCTGTTCCTCCCGCAAAAAAATGACAAAAATGTAATTAAGAGTGGCTG 4145
 Db 60246 TAGACAGTAGTAGTCTTGTGCTTGTGAGACATGTAGATTATCAACCAATTATTAAGAAAA 60305
 QY 4146 ATATGCTAGGCGAGTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCGAGGGG 4205
 Db 60306 ATATAGCAGGTGGAGTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCGAGGGGTG 60365

RESULT 14
 US-09-078-294-3
 ; Sequence 3, Application US/09078294
 ; Patent No. 6265211
 ; GENERAL INFORMATION:
 ; APPLICANT: Choo, Kong-Hong Andy
 ; APPLICANT: Du Sart, Desiree
 ; APPLICANT: Cancilla, Michael R.
 ; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
 ; FILE REFERENCE: Davies COL
 ; CURRENT APPLICATION NUMBER: US/09/078,294
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2003, 16:05:11 ; Search time 955 Seconds
(without alignments)
14611.055 Million cell updates/sec

Title: US-09-845-020A-5

Perfect score: 6235

Sequence: 1 gatcacttgagacagtagt.....tgcacttccacagcgt 6235

Scoring table:

IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6235	100.0	6235	11	US-09-845-020A-5
2	6235	100.0	6679	11	US-09-845-020A-1
3	2834	45.5	2834	11	US-09-845-020A-6
4	1252	20.1	1252	11	US-09-845-020A-7
5	1002.2	16.1	36221	11	US-09-954-556-29
6	744.8	11.9	260209	12	US-10-025-966A-23
7	744.8	11.9	260209	12	US-10-265-071-23
8	683.2	11.0	91000	14	US-10-002-491-10
9	670	10.7	104000	14	US-10-012-984-14
10	652.8	10.5	10351	9	US-09-874-470-5
11	647.2	10.4	161552	15	US-10-081-327-40
12	645.4	10.4	7126	12	US-10-017-161-1529
13	588.6	9.4	128779	15	US-10-081-327-38
14	583.8	9.4	6766	10	US-09-764-847-1878
15	583.8	9.4	6766	14	US-10-092-154-1878
16	564.4	9.1	28818	10	US-09-764-877-2266

17	563.8	9.0	9704	13	US-10-109-860-3	Sequence 3, Appl1
18	563.8	9.0	175590	11	US-09-911-077A-13	Sequence 13, Appl
19	562.4	9.0	54945	11	US-09-967-669-10	Sequence 10, Appl
20	562	9.0	2845	13	US-10-027-632-112134	Sequence 112134,
21	560	9.0	1503841	9	US-09-795-668-1	Sequence 1, Appl1
22	560	9.0	1503841	9	US-09-795-668-1	Sequence 1, Appl1
23	560	9.0	1503841	10	US-09-946-807-1	Sequence 1, Appl1
24	533.6	8.6	19172	10	US-09-764-877-3474	Sequence 3474, Ap
25	533	8.5	1192	13	US-10-027-632-118931	Sequence 118931,
26	522.2	8.4	5281	11	US-09-764-881-6949	Sequence 6949, Ap
27	509	8.2	2324	13	US-10-027-632-102474	Sequence 102474,
28	506.4	8.1	378361	11	US-09-901-136-3	Sequence 1256, Ap
29	500.2	8.0	776	10	US-09-764-847-1256	Sequence 1256, Ap
30	500.2	8.0	776	14	US-10-092-154-1256	Sequence 1257, Ap
31	500.2	8.0	900	10	US-09-764-847-1257	Sequence 1257, Ap
32	500.2	8.0	900	14	US-10-092-154-1257	Sequence 1257, Ap
33	499.6	8.0	142299	11	US-09-911-077A-14	Sequence 14, Appl
34	497.8	8.0	300000	14	US-10-262-552-33	Sequence 33, Appl
35	490.4	7.9	78025	14	US-10-020-141-9	Sequence 9, Appl1
36	486.4	7.8	27148	9	US-09-764-860-1046	Sequence 1046, Ap
37	486.4	7.8	27148	14	US-10-074-095-1046	Sequence 1046, Ap
38	468.6	7.5	3824	11	US-09-931-836-22	Sequence 22, Appl
39	468.6	7.5	3824	12	US-10-035-977-22	Sequence 22, Appl
40	468.6	7.5	3824	12	US-10-137-870-541	Sequence 541, App
41	468.6	7.5	3824	12	US-10-140-018-541	Sequence 541, App
42	468.6	7.5	3824	12	US-10-140-021-541	Sequence 541, App
43	468.6	7.5	3824	12	US-10-140-274-541	Sequence 541, App
44	468.6	7.5	3824	12	US-10-140-471-541	Sequence 541, App
45	468.6	7.5	3824	12	US-10-140-807-541	Sequence 541, App

ALIGNMENTS

RESULT 1
US-09-845-020A-5
Sequence 5, Application US/09845020A
Publication No. US20030022850A1
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heartlein, Michel W.
TITLE OF INVENTION: Selden, Richard F.
TITLE OF INVENTION: Genomic Sequences for Protein Production
FILE REFERENCE: 50010/017003
CURRENT FILING DATE: 2001-04-27
CURRENT APPLICATION NUMBER: US 09/845,020A
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: US 60/084,649
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 6235
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-020A-5
Query Match
Best Local Similarity 100.0%: Score 6235; DB 11; Length 6235;
Matches 6235: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCACTTGAGACAGTAGTTCACAGCAGCTGGGACATAGGAGAGACTGCTCTACG 60
DB 1 GATCACTTGAGACAGTAGTTCACAGCAGCTGGGACATAGGAGAGACTGCTCTACG 60
QY 61 AAAATTCAAAAATATGCGCGGCGATGCTGCTACGCTCTAATCCCTGAACCTTGGG 120
DB 61 AAAATTCAAAAATATGCGCGGCGATGCTGCTACGCTCTAATCCCTGAACCTTGGG 120
QY 121 ACATAGAGCAAGTGATGATGCTGAGTTCAGAGTTCGAGTACGCTGCGCAACATGCT 180
DB 121 ACATAGAGCAAGTGATGATGCTGAGTTCAGAGTTCGAGTACGCTGCGCAACATGCT 180

D	b		121	ACATCAGGCGAAGTGGATCACTTGAGGTGCAGAGATTGCAGACTACGCTGGCCAAACATGGT	180
O	y		181	GAAACCCATCTCCACTAAAAAATTACAATAATTAGCCAGCATGGTGGCAGGCACTGTA	240
D	b		181	GAAACCCATCTCCACTAAAAAATTACAATAATTAGCCAGCATGGTGGCAGGCACTGTA	240
O	y		241	ATCCGGCTACTCAGAGAGCTGAGGCAGGAGAATCACTTGAACCAGGAGCGGAGTTG	300
D	b		241	ATCCGGCTACTCAGAGAGCTGAGGCAGGAGAATCACTTGAACCAGGAGCGGAGTTG	300
O	y		301	CAGTAGCTGAGATCACACCACTGCACGCCCGGGGTGCACAGCAAAAGCTATATCTC	360
D	b		301	CAGTAGCTGAGATCACACCACTGCACGCCCGGGGTGCACAGCAAAAGCTATATCTC	360
O	y		361	AAAAAAAAATAAAAAATTTAGCCAGGCATGGTAGGCACACTCTAGTCAG	420
D	b		361	AAAAAAAAATAAAAAATTTAGCCAGGCATGGTAGGCACACTCTAGTCAG	420
O	y		421	CTACTCAGAGAGCTGAGGTGGGAGGATCACTTGAACCTGGGGCAGTCAGGCTACAGTGA	480
D	b		421	CTACTCAGAGAGCTGAGGTGGGAGGATCACTTGAACCTGGGGCAGTCAGGCTACAGTGA	480
O	y		481	GCCAAGATCATGCCACTCACTCCAGCCTGGGGCAACAGAGAGACCTGTCTCTAAA	540
D	b		481	GCCAAGATCATGCCACTCACTCCAGCCTGGGGCAACAGAGAGACCTGTCTCTAAA	540
O	y		541	AATAATATTAATAAAGAAAAAACAGCTCTGTTATGTCTCGTCCATACATCTACT	600
D	b		541	AATAATATTAATAAAGAAAAAACAGCTCTGTTATGTCTCGTCCATACATCTACT	600
O	y		601	ATGTATATAGTTTGCANAACCTCAAGATCCAGATAGTCAATTTTTAGGCTTGTGGCCGT	660
D	b		601	ATGTATATAGTTTGCANAACCTCAAGATCCAGATAGTCAATTTTTAGGCTTGTGGCCGT	660
O	y		661	ATGGCTCTGTGCACAATCACTCTGCCCTGTCTTCTTACACAAAAAGCATATAACAT	720
D	b		661	ATGGCTCTGTGCACAATCACTCTGCCCTGTCTTCTTACACAAAAAGCATATAACAT	720
O	y		721	ACATACATGAATTTTTATAGACATCGAGATTTGAATTCATATGATTTTTACATTTAT	780
D	b		721	ACATACATGAATTTTTATAGACATCGAGATTTGAATTCATATGATTTTTACATTTAT	780
O	y		781	AAAAATATCTTTTAAAAATTTTCCCCTAACCATTTAAAGTGTAAAGCCGGCCAGCGC	840
D	b		781	AAAAATATCTTTTAAAAATTTTCCCCTAACCATTTAAAGTGTAAAGCCGGCCAGCGC	840
O	y		841	GCCATCGTCAGCGCCGTATATCCAGCAATTTGGGAGGCTGAGGTGGCGAGATCACTTAG	900
D	b		841	GCCATCGTCAGCGCCGTATATCCAGCAATTTGGGAGGCTGAGGTGGCGAGATCACTTAG	900
O	y		901	ATCAACATGTTGSAAGACAGCCTGGCCAACTATGCAAAAACCCATTCTTACTTAAAAATTA	960
D	b		901	ATCAACATGTTGSAAGACAGCCTGGCCAACTATGCAAAAACCCATTCTTACTTAAAAATTA	960
O	y		961	AAAAATAGCTGGGCACTAGTGTGTCACACCTGTGATCCAGCTACTTGGGAGCTGAGGCA	1020
D	b		961	AAAAATAGCTGGGCACTAGTGTGTCACACCTGTGATCCAGCTACTTGGGAGCTGAGGCA	1020
O	y		1021	GGAGAAATGCGTTGAACCTGGGAAGGGAGTTGCTGATGAGCCACATATGCCCCACT	1080
D	b		1021	GGAGAAATGCGTTGAACCTGGGAAGGGAGTTGCTGATGAGCCACATATGCCCCACT	1080
O	y		1081	TCCAGCCTGGGGGACAGAGTGAAGATTTGCTCACAAGAAAAAAAATGTAAAAAGCAT	1140
D	b		1081	TCCAGCCTGGGGGACAGAGTGAAGATTTGCTCACAAGAAAAAAAATGTAAAAAGCAT	1140
O	y		1141	TCCTAATTCAGTGTACATCACTGATCACTACAGGCTGCGTACTCCTGCTGAGGCAT	1200
D	b		1141	TCCTAATTCAGTGTACATCACTGATCACTACAGGCTGCGTACTCCTGCTGAGGCAT	1200
O	y		1201	ACCTGAGAAGTAGAGTTCGTTGGTGCACAGGACATACATTTCCACATTAACATAGACCT	1260
D	b		1201	ACCTGAGAAGTAGAGTTCGTTGGTGCACAGGACATACATTTCCACATTAACATAGACCT	1260

QY	1261	ACCAAGTTGCATCCAGAGAGGTTTTTTTTTTTACAAATCTACACTCCGCCAGCAACAAT	1320
DB	1261	ACCAAGTTGCATCCAGAGAGGTTTTTTTTTTTACAAATCTACACTCCGCCAGCAACAAT	1320
QY	1321	GAGAGTTACTCCAGATCCTTTACAAAGATGCTCTAAGCCAGTACGATGAAACAGGA	1380
DB	1321	GAGAGTTACTCCAGATCCTTTACAAAGATGCTCTAAGCCAGTACGATGAAACAGGA	1380
QY	1381	AGTGGAGGGAGAGCTGCCAGCCCTCTTAACATGGAAGAAATACCTGGTAGACCTTCT	1440
DB	1381	AGTGGAGGGAGAGCTGCCAGCCCTCTTAACATGGAAGAAATACCTGGTAGACCTTCT	1440
QY	1441	GGATGCTGGAAGATGAAATAACGGGGGGTCTCTGAGAGCTGGCCCCCTCTCAGATCACTGTG	1500
DB	1441	GGATGCTGGAAGATGAAATAACGGGGGGTCTCTGAGAGCTGGCCCCCTCTCAGATCACTGTG	1500
QY	1501	ACTTCTGAGCCTCAAGTCCAGTCTCAGCCCCATGTCTATGAGCCAGATTAATGAGCCT	1560
DB	1501	ACTTCTGAGCCTCAAGTCCAGTCTCAGCCCCATGTCTATGAGCCAGATTAATGAGCCT	1560
QY	1561	CACCTCTGTTGGTCTTTATTTCCCCATGTGGGGGTGAAGTCTGGATTGAGCCGTAT	1620
DB	1561	CACCTCTGTTGGTCTTTATTTCCCCATGTGGGGGTGAAGTCTGGATTGAGCCGTAT	1620
QY	1621	TCAAGATGTACAGCTTCTTGACAGAAAGTAGTCTACAGAAACACAGAGGGCTTGCA	1680
DB	1621	TCAAGATGTACAGCTTCTTGACAGAAAGTAGTCTACAGAAACACAGAGGGCTTGCA	1680
QY	1681	AGATGATCTAATCTGCAAATCCTACCTGGCTCAGCCACAGCTAGTTCTGTATCTTAAC	1740
DB	1681	AGATGATCTAATCTGCAAATCCTACCTGGCTCAGCCACAGCTAGTTCTGTATCTTAAC	1740
QY	1741	AAGTTTTTCACTTCTCTGAGAGGCATCCCTGGCTACAAACACACATGGTTGACAGGA	1800
DB	1741	AAGTTTTTCACTTCTCTGAGAGGCATCCCTGGCTACAAACACACATGGTTGACAGGA	1800
QY	1801	TGAATATGACGAAGTCCCTTACACCTGTATATCCACAGCACTTGGGAGGACCAAGCGGGGTGG	1860
DB	1801	TGAATATGACGAAGTCCCTTACACCTGTATATCCACAGCACTTGGGAGGACCAAGCGGGGTGG	1860
QY	1861	ATGGCTTGAGCCTGAGAGGTACAGACATGCCGGCAGTCTCAGAGCCCTGTGCTCTC	1920
DB	1861	ATGGCTTGAGCCTGAGAGGTACAGACATGCCGGCAGTCTCAGAGCCCTGTGCTCTC	1920
QY	1921	GGGGCCTCCTGCTGGGCTCCCACTTCCGCTGGGCACTTGAAGAGCCCTTACGCCACCG	1980
DB	1921	GGGGCCTCCTGCTGGGCTCCCACTTCCGCTGGGCACTTGAAGAGCCCTTACGCCACCG	1980
QY	1981	CTGCACGTGGGAGCCCTTCTTGAGGGTGGCCAGAGCCAGCGGCTCCCTCAGCTTGC	2040
DB	1981	CTGCACGTGGGAGCCCTTCTTGAGGGTGGCCAGAGCCAGCGGCTCCCTCAGCTTGC	2040
QY	2041	AGGAGGTGTGAGAGGAGAGGCTCAACAGAGAAACGGGGGCTGCGCACGGGCTTGCGGGC	2100
DB	2041	AGGAGGTGTGAGAGGAGAGGCTCAACAGAGAAACGGGGGCTGCGCACGGGCTTGCGGGC	2100
QY	2101	CAGCTGAGATTCGGGGTGGGGGTGGGCTTGGCGGGGCCCGACACTCTGGAGACAGCGGGCAG	2160
DB	2101	CAGCTGAGATTCGGGGTGGGGGTGGGCTTGGCGGGGCCCGACACTCTGGAGACAGCGGGCAG	2160
QY	2161	CCCTGCGCAGGCGCCGGGCAATGAAGAGCTTATGACACCGGGGCAACGGGCTCGGGAGGGTGT	2220
DB	2161	CCCTGCGCAGGCGCCGGGCAATGAAGAGCTTATGACACCGGGGCAACGGGCTCGGGAGGGTGT	2220
QY	2221	ACTGGGTGCCCCAGACAGTGCACAGCCCGCGCGGCTGTGCTCGATTTCTCAGTGGC	2280
DB	2221	ACTGGGTGCCCCAGACAGTGCACAGCCCGCGCGGCTGTGCTCGATTTCTCAGTGGC	2280
QY	2281	CTTAGCAGACCTTCCCGGGGGCAGGGCTCGGGAACCTGACAGCCGCAATGCTTGAAGCTTC	2340
DB	2281	CTTAGCAGACCTTCCCGGGGGCAGGGCTCGGGAACCTGACAGCCGCAATGCTTGAAGCTTC	2340

QY 2341 CCTCATAGGCTCTGTGGGCGCGAGCCTCCCGACAGCACCCTCTGTCCACAG 2400
|||||
Db 2341 CCTCATAGGCTCTGTGGGCGCGAGCCTCCCGACAGCACCCTCTGTCCACAG 2400
QY 2401 CGCCAGCTCCCATGACACGCAAGGCTGAGAAATGGGGGCGACAGGCGGAGCTGG 2460
|||||
Db 2401 CGCCAGCTCCCATGACACGCAAGGCTGAGAAATGGGGGCGACAGGCGGAGCTGG 2460
QY 2461 CAGGAGCTACCCCTGACAGCCCTGTGGGAAATCCACTGGGTGAAGCCAGTGGCTCT 2520
|||||
Db 2461 CAGGAGCTACCCCTGACAGCCCTGTGGGAAATCCACTGGGTGAAGCCAGTGGCTCT 2520
QY 2521 GAGTCTGGTGGAGACTTGGAAACCTTATGTCTAGCTCAGGAGTCTGAATTAACCAAT 2580
|||||
Db 2521 GAGTCTGGTGGAGACTTGGAAACCTTATGTCTAGCTCAGGAGTCTGAATTAACCAAT 2580
QY 2581 CAGACCCCTGTGTAGCTCAGGCTGTGTGAATGCAACCAATCCACTCTGTATCTAGCT 2640
|||||
Db 2581 CAGACCCCTGTGTAGCTCAGGCTGTGTGAATGCAACCAATCCACTCTGTATCTAGCT 2640
QY 2641 ACTGTGATGGGCGCTTGGAGAACCTTATGTCTAGCTCAGGAGTCTGAATTAACCAATC 2700
|||||
Db 2641 ACTGTGATGGGCGCTTGGAGAACCTTATGTCTAGCTCAGGAGTCTGAATTAACCAATC 2700
QY 2701 GGCACCTGTATCTAGCTCAAGTTTGTAAACACACCAATCAGACCCCTGTGTAGCTC 2760
|||||
Db 2701 GGCACCTGTATCTAGCTCAAGTTTGTAAACACACCAATCAGACCCCTGTGTAGCTC 2760
QY 2761 AGGCTATGTGAATGACCAATCGACAGTCTGTATCTGTGGCTACTTTCATGGCCATCCGTGT 2820
|||||
Db 2761 AGGCTATGTGAATGACCAATCGACAGTCTGTATCTGTGGCTACTTTCATGGCCATCCGTGT 2820
QY 2821 GAAGAGACCCACCAAGGCTTGTGTGAGCAATTAAGCTCTATACCTGGGTGGAGCT 2880
|||||
Db 2821 GAAGAGACCCACCAAGGCTTGTGTGAGCAATTAAGCTCTATACCTGGGTGGAGCT 2880
QY 2881 GGGCTGAGTCCGAAAGAGAGTCAAGGAGAGATAGGCTGGGCGCTTGTATAGAT 2940
|||||
Db 2881 GGGCTGAGTCCGAAAGAGAGTCAAGGAGAGATAGGCTGGGCGCTTGTATAGAT 2940
QY 2941 TTGGGTAGGTAAAGAAATTTACAGTCAAGGGGCTTGTCTCTGGCGGCGAGAGTGG 3000
|||||
Db 2941 TTGGGTAGGTAAAGAAATTTACAGTCAAGGGGCTTGTCTCTGGCGGCGAGAGTGG 3000
QY 3001 GGGGTGCAAGGTCTGAGTGGGGGTCTTTTGAAGCAGAGATGAGCCAGAAAGAGCT 3060
|||||
Db 3001 GGGGTGCAAGGTCTGAGTGGGGGTCTTTTGAAGCAGAGATGAGCCAGAAAGAGCT 3060
QY 3061 TTCAACAAGTAAATGTCAATTAAGGCAAGGACCCGCAATTAACCTCTTTTGGGTG 3120
|||||
Db 3061 TTCAACAAGTAAATGTCAATTAAGGCAAGGACCCGCAATTAACCTCTTTTGGGTG 3120
QY 3121 GAATGTCAATCAGTTAACTTGGGGGAGGGCATATTAATCACTCTTTTGTGATTTCAAGTTAC 3180
|||||
Db 3121 GAATGTCAATCAGTTAACTTGGGGGAGGGCATATTAATCACTCTTTTGTGATTTCAAGTTAC 3180
QY 3181 TTTCAGGCAATCTGGGCGGTATATGTGCAAGTTACAGGGGATGCGATGGCTTGGCT 3240
|||||
Db 3181 TTTCAGGCAATCTGGGCGGTATATGTGCAAGTTACAGGGGATGCGATGGCTTGGCT 3240
QY 3241 CAGAGGCTTGACACTACTCTGTGTGGGCTTGGAGAAATTTTGTGTGACACTGTAT 3300
|||||
Db 3241 CAGAGGCTTGACACTACTCTGTGTGGGCTTGGAGAAATTTTGTGTGACACTGTAT 3300
QY 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTGTGTCTAGCTCAGGGATTTGAAGCA 3360
|||||
Db 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTGTGTCTAGCTCAGGGATTTGAAGCA 3360
QY 3361 CCATATCAGCCCTGTGAAAAACACACTGGCTCTACCAATCAGCAGGATGTGGGTGG 3420
|||||
Db 3361 CCATATCAGCCCTGTGAAAAACACACTGGCTCTACCAATCAGCAGGATGTGGGTGG 3420
QY 3421 GGCAGATTAAGAAATTAAGCAGGCTGGCCGAGCCAGCAATGGCAACGCCACAGGCTC 3480
|||||

Db 3421 GGCAGATTAAGAAATTAAGCAGGCTGGCCGAGCCAGCAATGGCAACGCCACAGGCTC 3480
|||||
QY 3481 CTATCCACAAATATGAGCAGCTTGTCTTTTGTCTGTGGTGAATTAATCTGTACTGTG 3540
|||||
Db 3481 CTATCCACAAATATGAGCAGCTTGTCTTTTGTCTGTGGTGAATTAATCTGTACTGTG 3540
QY 3541 CTTTTTGGTCCACACCTCTTTTATGAGCTGTAACTACACAGGAAGCTGTGACGCTTC 3600
|||||
Db 3541 CTTTTTGGTCCACACCTCTTTTATGAGCTGTAACTACACAGGAAGCTGTGACGCTTC 3600
QY 3601 ACTCCTGAAGCCACTTAAGACACAGAGCCACCGGAGGAATGAACAACTCCGGCGGCT 3660
|||||
Db 3601 ACTCCTGAAGCCACTTAAGACACAGAGCCACCGGAGGAATGAACAACTCCGGCGGCT 3660
QY 3661 GCCTTAAGAGCTATTAACCTACCGCGAAGTCTGACCTTCACTCTCAGCAGGAGA 3720
|||||
Db 3661 GCCTTAAGAGCTATTAACCTACCGCGAAGTCTGACCTTCACTCTCAGCAGGAGA 3720
QY 3721 CCAGAACCCACAGAAAGAAACCTGCAACATCTGAACATCAGAAAGCAAACT 3780
|||||
Db 3721 CCAGAACCCACAGAAAGAAACCTGCAACATCTGAACATCAGAAAGCAAACT 3780
QY 3781 CCAGATGCAACACCTTAAGAGCTGTAACTACACTGCGAGGCTCCGCTTCTTCTTG 3840
|||||
Db 3781 CCAGATGCAACACCTTAAGAGCTGTAACTACACTGCGAGGCTCCGCTTCTTCTTG 3840
QY 3841 AAGTCAATGAGAACCAAGCACTACCACTTTGGGACACAAAGCCAGAGTTGATCAGC 3900
|||||
Db 3841 AAGTCAATGAGAACCAAGCACTACCACTTTGGGACACAAAGCCAGAGTTGATCAGC 3900
QY 3901 CTGGGCAACATGATGAATGACCTCTCTGCAAAAAAAATTAACAAAAATTTGGGG 3960
|||||
Db 3901 CTGGGCAACATGATGAATGACCTCTCTGCAAAAAAAATTAACAAAAATTTGGGG 3960
QY 3961 AGCATGTGTGTCCTGCTGCTGTGCTCCAGCTACCGGGGAGGCTAAAGTGGAGATCTCT 4020
|||||
Db 3961 AGCATGTGTGTCCTGCTGCTGTGCTCCAGCTACCGGGGAGGCTAAAGTGGAGATCTCT 4020
QY 4021 TGAGCTTGGAGGTGAACACTGACAGTGTGATTTTACCACAGCCCTTAGCTGG 4080
|||||
Db 4021 TGAGCTTGGAGGTGAACACTGACAGTGTGATTTTACCACAGCCCTTAGCTGG 4080
QY 4081 GGACAGACTGAGACCTGTTCCCTCGCAAAAAATTAACAAAAATTTGAAGTATAGAGT 4140
|||||
Db 4081 GGACAGACTGAGACCTGTTCCCTCGCAAAAAATTAACAAAAATTTGAAGTATAGAGT 4140
QY 4141 GCTGTATATGCTAGGCGCAAGTGGCTCAATGCTTAATCCAGCACTTGGGAAGCCGAG 4200
|||||
Db 4141 GCTGTATATGCTAGGCGCAAGTGGCTCAATGCTTAATCCAGCACTTGGGAAGCCGAG 4200
QY 4201 GCGGGCGGCTACCTTAAGGTCAAGAGTGTGAGACAGGCTGGCAACATGGAAGAGCC 4260
|||||
Db 4201 GCGGGCGGCTACCTTAAGGTCAAGAGTGTGAGACAGGCTGGCAACATGGAAGAGCC 4260
QY 4261 ATCTCTTTAAAAAATTAACAAATTAAGCCGCTGTGGGGCACTGTGGAGCATGCTGTAA 4320
|||||
Db 4261 ATCTCTTTAAAAAATTAACAAATTAAGCCGCTGTGGGGCACTGTGGAGCATGCTGTAA 4320
QY 4321 TCCAGCTACTCAGAGGCTGAGGAGAGAAATCACTGGAACCCAGAGGGGCGGCTTGC 4380
|||||
Db 4321 TCCAGCTACTCAGAGGCTGAGGAGAGAAATCACTGGAACCCAGAGGGGCGGCTTGC 4380
QY 4381 AGTGAAGCAGAGATGTGCATTTGCACTCCACCACTCAGGCTGGGCAACAGAGCCAAA 4440
|||||
Db 4381 AGTGAAGCAGAGATGTGCATTTGCACTCCACCACTCAGGCTGGGCAACAGAGCCAAA 4440
QY 4441 CTCTGTCTTAAAAAATTAAGAGTGTGACATTTAAGAGTGTGCAATGCAATAG 4500
|||||
Db 4441 CTCTGTCTTAAAAAATTAAGAGTGTGACATTTAAGAGTGTGCAATGCAATAG 4500
QY 4501 TTGCAGGCAACATGTTAAATGTGAGTCTGCTGCTTCATAGTCTCTGTAAAAAAC 4560
|||||

Dh 20 GATCATTGAGGACAGTAGTTCAGAACCCAGCCTGGGACATAGGAGACTGTCTACG 79
Oy 61 AAAATCAAAAAATATGCGCGGATGTTGGCTACCTCTATCCCTGACCTTGGG 120
Dh 80 AAAATTCAAAAATATGCGCGGATGTTGGCTACCTCTATCCCTGACCTTGGG 139
Oy 121 ACATCAAGGCAAGTAGATCATTGAGTTCAGAGTTCAGACTAGCTGGCCAMCATGT 180
Dh 140 ACATCAAGGCAAGTAGATCATTGAGTTCAGAGTTCAGACTAGCTGGCCAMCATGT 199
Oy 181 GAAACCTTATCTCCACTAAAAATTCAGAGGATGAGGATGAGGAGGAGGAGGAGG 240
Dh 200 GAAACCTTATCTCCACTAAAAATTCAGAGGATGAGGATGAGGAGGAGGAGGAGG 259
Oy 241 ATCCGGCTACTGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGG 300
Dh 260 ATCCGGCTACTGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGG 319
Oy 301 CAGTGAAGTGAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
Dh 320 CAGTGAAGTGAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 379
Oy 361 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
Dh 380 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 439
Oy 421 CTACTCAGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAG 480
Dh 440 CTACTCAGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAG 499
Oy 481 GCCAAGATCATGCCACTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Dh 500 GCCAAGATCATGCCACTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 559
Oy 541 AATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
Dh 560 AATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 619
Oy 601 ATGATATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGATAT 660
Dh 620 ATGATATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGATAT 679
Oy 661 ATGATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGATATGAT 720
Dh 680 ATGATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGATATGAT 739
Oy 721 ACATCAATGATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGAT 780
Dh 740 ACATCAATGATATGTTGCAAACTCAAGATCAGATATGATATGATATGATATGAT 799
Oy 781 AAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Dh 800 AAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 859
Oy 841 GCCATGCTGACGCTGTAATTCAGACTTGGAGGCTGAGTGGGAGGAGGAGGAGG 900
Dh 860 GCCATGCTGACGCTGTAATTCAGACTTGGAGGCTGAGTGGGAGGAGGAGGAGG 919
Oy 901 ATCAACAGTTCGAGACGAGCTGGCAACATAGCAAAACCCCATTTCTACTAAAA 960
Dh 920 ATCAACAGTTCGAGACGAGCTGGCAACATAGCAAAACCCCATTTCTACTAAAA 979
Oy 961 AAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1020
Dh 980 AAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1039
Oy 1021 GAGAAATGCTTGAACCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Dh 1040 GAGAAATGCTTGAACCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
Oy 1081 TCCAGCCTGGGTGACAGAGTGAAGCTTCTCAACGAAAAAAGTGTAAAAAGCAT 1140
Dh 1100 TCCAGCCTGGGTGACAGAGTGAAGCTTCTCAACGAAAAAAGTGTAAAAAGCAT 1159

Oy 1141 TCCTAATTCAGTACATCAGTACATCAGTACATCAGTACATCAGTACATCAGTAC 1200
Dh 1160 TCCTAATTCAGTACATCAGTACATCAGTACATCAGTACATCAGTACATCAGTAC 1219
Oy 1201 ACCTGAAGATGAGTTCGTTGTCAGAGGATCAGATCAGATCAGATCAGATCAGAT 1260
Dh 1220 ACCTGAAGATGAGTTCGTTGTCAGAGGATCAGATCAGATCAGATCAGATCAGAT 1279
Oy 1261 ACCAAGTTCAGTACAGAGGATTTTTTTTACAAATCTACACTCCCCCAGCAAAAT 1320
Dh 1280 ACCAAGTTCAGTACAGAGGATTTTTTTTACAAATCTACACTCCCCCAGCAAAAT 1339
Oy 1321 GAGATTCCTCAGATCCTTTTCAAAAGATGCTTAAGCCATACAGATGAAAAAG 1380
Dh 1340 GAGATTCCTCAGATCCTTTTCAAAAGATGCTTAAGCCATACAGATGAAAAAG 1399
Oy 1381 AGTGGAGGAGGAGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Dh 1400 AGTGGAGGAGGAGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1459
Oy 1441 GATGCTGGAAGATGATTAACGGGGGCTCTGAGAGCTGCCCCCTGTACATCACTG 1500
Dh 1460 GATGCTGGAAGATGATTAACGGGGGCTCTGAGAGCTGCCCCCTGTACATCACTG 1519
Oy 1501 ACTTGAGAGCTTCAGATCAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Dh 1520 ACTTGAGAGCTTCAGATCAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1579
Oy 1561 CACTCTGTTGTTGTTTATCTCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Dh 1580 CACTCTGTTGTTGTTTATCTCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1639
Oy 1621 TCAAGATGACAGCTTTCTTGACAGAAAGTGTGTACAGAAACAGCAGGAGGCTT 1680
Dh 1640 TCAAGATGACAGCTTTCTTGACAGAAAGTGTGTACAGAAACAGCAGGAGGCTT 1699
Oy 1681 AGATGATTCATGCAAAATCCTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
Dh 1700 AGATGATTCATGCAAAATCCTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1759
Oy 1741 AAGTTTTCACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
Dh 1760 AAGTTTTCACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1819
Oy 1801 TGAATGACAGAGTCCCTTACACCTGTAATCCAGACATTTGGAGGAGGAGGAGG 1860
Dh 1820 TGAATGACAGAGTCCCTTACACCTGTAATCCAGACATTTGGAGGAGGAGGAGG 1879
Oy 1861 ATGGCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Dh 1880 ATGGCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1939
Oy 1921 GGGGCTCTGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
Dh 1940 GGGGCTCTGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1999
Oy 1981 CTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
Dh 2000 CTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2059
Oy 2041 AGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
Dh 2060 AGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2119
Oy 2101 CAGCTGAGATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Dh 2120 CAGCTGAGATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2179
Oy 2161 CCTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Dh 2180 CCTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2239

QY 2221 ACTGGGCCCCAGCACTGGCAGGCCCGGCGCTGCTGCTGCATTTCTCACTGGGC 2280
Db 2240 ACTGGGGCCCCACACACTGGCCAGCCCGCGCGCTGCTGCTGCATTTCTCACTGGGC 2299
QY 2281 CTTAGACACCTTCCCGGGGAGAGGGCTGGGACCTTCAGGCCGSCATGCTGAGCCTCC 2340
Db 2300 CTTAGACACCTTCCCGGGGAGAGGGCTGGGACCTTCAGGCCGSCATGCTGAGCCTCC 2359
QY 2341 CTTCCATGGGCTCTGTGGGCCCGAGCCTCCCGAGCAGACCAACCCCTGCTGCAGAG 2400
Db 2360 CTTCCATGGGCTCTGTGGGCCCGAGCCTCCCGAGCAGACCAACCCCTGCTGCAGAG 2419
QY 2401 CGCCCATGTCATGAGCAGCAGAGGGCTGAGAGTGGCGGGGAGACCCGGGACTGG 2460
Db 2420 CGCCCATGTCATGAGCAGCAGAGGGCTGAGAGTGGCGGGGAGACCCGGGACTGG 2479
QY 2461 CAGGACACTACCCCTGCAGCCCTGTGTGCGAGATCCAGTGGGTGAAGCAGCTGGCTCT 2520
Db 2480 CAGGACACTACCCCTGCAGCCCTGTGTGCGAGATCCAGTGGGTGAAGCAGCTGGCTCT 2539
QY 2521 GAGTCTGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCGTAAATACCAAT 2580
Db 2540 GAGTCTGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCGTAAATACCAAT 2599
QY 2581 CAGCAGCCTGTGTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTATCTAGCT 2640
Db 2600 CAGCAGCCTGTGTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTGTATCTAGCT 2659
QY 2641 ACTGTAGGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATTTGAATACCAATC 2700
Db 2660 ACTGTAGGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATTTGAATACCAATC 2719
QY 2701 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTCTAGCTC 2760
Db 2720 GGCACCTGTATCTAGCTCAAGGTTTGTAAACACCAATCAGACCCCTGTCTAGCTC 2779
QY 2761 AGGCTATGTAAATGACCAATCGACCTGTATCTAGCTCAGCTTCAATGGGCACTCCGT 2820
Db 2780 AGGCTATGTAAATGACCAATCGACCTGTATCTAGCTCAGCTTCAATGGGCACTCCGT 2839
QY 2821 GAAGAGACCAACAAAGGCTTGTGTAGCAATTAAGCTTCTATCACTGGGTCAGAGT 2880
Db 2840 GAAGAGACCAACAAAGGCTTGTGTAGCAATTAAGCTTCTATCACTGGGTCAGAGT 2899
QY 2881 GGGCTGAGTCCGAAAGAGAGTCAAGCAGGAGATTAAGGTGGGCGCTTTTATAGAT 2940
Db 2900 GGGCTGAGTCCGAAAGAGAGTCAAGCAGGAGATTAAGGTGGGCGCTTTTATAGAT 2959
QY 2941 TTGGCTAGTAAAGGAAATTTACATCAAGGCGGCTTGTCTCTGGCGGGCAGAGTGG 3000
Db 2960 TTGGCTAGTAAAGGAAATTTACATCAAGGCGGCTTGTCTCTGGCGGGCAGAGTGG 3019
QY 3001 GGGGTGCAGAGTGTCTAGTGGGGGCTTTTGTAGCCAGAGTACGAGGAAAGAGACT 3060
Db 3020 GGGGTGCAGAGTGTCTAGTGGGGGCTTTTGTAGCCAGAGTACGAGGAAAGAGACT 3079
QY 3061 TTTCACAAAGTAAATGTCAATTAAGGCAAGAGACCCGCATTTACACCTTTTGTGTG 3120
Db 3080 TTTCACAAAGTAAATGTCAATTAAGGCAAGAGACCCGCATTTACACCTTTTGTGTG 3139
QY 3121 GAATGTCAATAGTTAAGTTGGGGCAGGSCATTTCACTTTTGTGTATCTTCACTTAC 3180
Db 3140 GAATGTCAATAGTTAAGTTGGGGCAGGSCATTTCACTTTTGTGTATCTTCACTTAC 3199
QY 3181 TTTCAGGCACTGTGGGCTATATGTGCAAGTTACAGGAGATCGATGGCTTGGCTGGCT 3240
Db 3200 TTTCAGGCACTGTGGGCTATATGTGCAAGTTACAGGAGATCGATGGCTTGGCTGGCT 3259
QY 3241 CAAAGGCTTACAGCTACTGTGGTGGGGCTTGGAGATGTTGTGTGCACTCTGTAT 3300
Db 3260 CAAAGGCTTACAGCTACTGTGGTGGGGCTTGGAGATGTTGTGTGCACTCTGTAT 3319
QY 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTTGTGTCTAGCTCAGGAGATTTGTAACGCA 3360

Db 3320 CTTAGTTAATCTAGTGGGAGCTGGAGAACCTTTGTGTCTAGCTCAGGAGATTTGTAACGCA 3379
QY 3361 CCAATTCAGCGCCCTGTCAAAAACAGACACTGCGCTTACCAATTCAGAGATGTGGTGG 3420
Db 3380 CCAATTCAGCGCCCTGTCAAAAACAGACACTGCGCTTACCAATTCAGAGATGTGGTGG 3439
QY 3421 GGGCAGATTAAGAAATTAAGAGCAGCTGCCGAGCAGCAGTGGCAACGGCACAGCTCC 3480
Db 3440 GGGCAGATTAAGAAATTAAGAGCAGCTGCCGAGCAGCAGTGGCAACGGCACAGCTCC 3499
QY 3481 CTATCCAAATATGAGCCTTGTCTTGTGTGTGTGCAATTAATCTGCTACTGCTCG 3540
Db 3500 CTATCCAAATATGAGCCTTGTCTTGTGTGTGTGCAATTAATCTGCTACTGCTCG 3559
QY 3541 CTTTGTGGGTCCACACTGCTTTTATGAGCTGTAACTCACTCAACAGAGTCTGCAGCTTC 3600
Db 3560 CTTTGTGGGTCCACACTGCTTTTATGAGCTGTAACTCACTCAACAGAGTCTGCAGCTTC 3619
QY 3601 ACTCCTAAGGCACTAAGACACGAGCCACCGGAGGAATGAACAATCCCGCGCGCT 3660
Db 3620 ACTCCTAAGGCACTAAGACACGAGCCACCGGAGGAATGAACAATCCCGCGCGCT 3679
QY 3661 GCCTTAAGAGCTATAACACTCACCGGAGAGTCTGCAGCTTCACTCTCAGCCGCGAGA 3720
Db 3680 GCCTTAAGAGCTATAACACTCACCGGAGAGTCTGCAGCTTCACTCTCAGCCGCGAGA 3739
QY 3721 CCACGAACCCACAGAGGAAGAAATGCGAACACATCTGAATCAAGAAAGAACAACT 3780
Db 3740 CCACGAACCCACAGAGGAAGAAATGCGAACACATCTGAATCAAGAAAGAACAACT 3799
QY 3781 CCAGATCACACCTTAAGAGCTGTAACTCACTCACTGCGAGGGTCCGCGCTTCTCTTG 3840
Db 3800 CCAGATCACACCTTAAGAGCTGTAACTCACTCACTGCGAGGGTCCGCGCTTCTCTTG 3859
QY 3841 AAGTCAATGAGACCAACACACTCAACAGTTTCCGACACAAGCCCGAGAGTTGAGATCAGC 3900
Db 3860 AAGTCAATGAGACCAACACACTCAACAGTTTCCGACACAAGCCCGAGAGTTGAGATCAGC 3919
QY 3901 CTTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTCAAAAAATTTGGCGG 3960
Db 3920 CTTGGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTCAAAAAATTTGGCGG 3979
QY 3961 AGCATGTGCTGCTGCTGTGTGTCCAGCTACGCGGAGGCTTAAGTGGGAGATGCT 4020
Db 3980 AGCATGTGCTGCTGCTGTGTGTCCAGCTACGCGGAGGCTTAAGTGGGAGATGCT 4039
QY 4021 TGAGCCTGGGAGGTGAAGACTGCAGTGTGATTTGTACCAACAGCCCTCTAGCTGGG 4080
Db 4040 TGAGCCTGGGAGGTGAAGACTGCAGTGTGATTTGTACCAACAGCCCTCTAGCTGGG 4099
QY 4081 GGACAGACTGAGACCCGTTTCCCTCCGCAAAAAAATTTGCAAAAAATGTAAAGAGT 4140
Db 4100 GGACAGACTGAGACCCGTTTCCCTCCGCAAAAAAATTTGCAAAAAATGTAAAGAGT 4159
QY 4141 GCTGTATATGCTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCCGAG 4200
Db 4160 GCTGTATATGCTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCCGAG 4219
QY 4201 GCGGGCGGGCACCTAAGGTACAGAGTGTAGACACCCGCGCAACATGAGCAAAAGCCC 4260
Db 4220 GCGGGCGGGCACCTAAGGTACAGAGTGTAGACACCCGCGCAACATGAGCAAAAGCCC 4279
QY 4261 ATCTCTTCTAAAAATATACAAATTAAGCGGCTGTGGGGCAGTGGTGGAGCATGCTGTAA 4320
Db 4280 ATCTCTTCTAAAAATATACAAATTAAGCGGCTGTGGGGCAGTGGTGGAGCATGCTGTAA 4339
QY 4321 TCCAGCTACTCAGAGGCTGTAGGCGAGAGATCACTTGAACCCAGAGGCGCGGTTGC 4380
Db 4340 TCCAGCTACTCAGAGGCTGTAGGCGAGAGATCACTTGAACCCAGAGGCGCGGTTGC 4399
QY 4381 AGTGAAGCGAGATCGTGCATTTGCACTCCACCCACTCCAGCCGCGGCAACAAAGCCAAA 4440

Db	4400	AGTGGCCGAGATGCTGGCATTGGACATCCACCACCTCCAGCTGGGCAACAAGGCCAA	4459
QY	4441	CTGTGTCTTAAAAAAAAAAAAAAAAAGTGCCTGACATTTAAGAGGTGTGCATGCAATAG	4500
Db	4460	CTGTGTCTTAAAAAAAAAAAAAAAAAGGCTGTCACATATAAGAGGTGTGCATGCAATAG	4519
QY	4501	TTGGCAGGCAACATGTTTAAAGAATGTGGAGTCTCGCTTCCTCCATGTCCTGTAAAAAC	4560
Db	4520	TTGGCAGGCAACATGTTTAAAGAATGTGGAGTCTCGCTTCCTCCATGTCCTGTAAAAAC	4579
QY	4561	CACCCCTCAAGGCCAGGTGTCAGTGGCTCATGGCTTAAATCCAGACACTTTGGAGGCCGAG	4620
Db	4580	CACCCCTCAAGGCCAGGTGTCAGTGGCTCATGGCTTAAATCCAGACACTTTGGAGGCCGAG	4639
QY	4621	GGGGGTGATCACTGAGTGACGATGAGATTGAGACACGCTTGACCAACAATGGTGAAT	4680
Db	4640	GGGGGTGATCACTGAGTGACGATGAGATTGAGACACGCTTGACCAACAATGGTGAAT	4699
QY	4681	CCCACTCTACTAAAAATACAAAATTTGATGAGATGGTGGTGCATGGCTATATCCAC	4740
Db	4700	CCCACTCTACTAAAAATACAAAATTTGATGAGATGGTGGTGCATGGCTATATCCAC	4759
QY	4741	CTACTTGGGAGCTGAGGCAAGAAATCCTAGAACACAGGGAGGCGAGGTGTAGTAGAG	4800
Db	4760	CTACTTGGGAGCTGAGGCAAGAAATCCTAGAACACAGGGAGGCGAGGTGTAGTAGAG	4819
QY	4801	CCGAGATCTGTCCCATTTGACATCCAGCCTGAGCAATGAGCGAACTCCATCTCAAAAAAC	4860
Db	4820	CCGAGATCTGTCCCATTTGACATCCAGCCTGAGCAATGAGCGAACTCCATCTCAAAAAAC	4879
QY	4861	AACAACAAAAAACCACCTCTCTACCTCCACGGAGGCTGGGTACAGAGCTGGGCGACATCAGT	4920
Db	4880	AACAACAAAAAACCACCTCTCTACCTCCACGGAGGCTGGGTACAGAGCTGGGCGACATCAGT	4939
QY	4921	GCAAGTGTCTGAGGCGACAGAGCTAAAGCGGAGTGTGACAGACCGCGGACACATATAACGTG	4980
Db	4940	GCAAGTGTCTGAGGCGACAGAGCTAAAGCGGAGTGTGACAGACCGCGGAGACATATAACGTG	4999
QY	4981	TGTGAGATCACTGTGTGAGATCAGACGTCCTGCCATTTGGTACACCACGGGGCCCCCA	5040
Db	5000	TGTGAGATCACTGTGTGAGATCAGACGTCCTGCCATTTGGTACACCACGGGGCCCCCA	5059
QY	5041	AGCACCGAGATGGCCCCCATCCAGTCCACCCATCCACTTCATCCAGAGATGTCGTGTT	5100
Db	5060	AGCACCGAGATGGCCCCCATCCAGTCCACCCATCCACTTCATCCAGAGATGTCGTGTT	5119
QY	5101	CTTGGCAGCGCTGGGGTAAATTAGAGACAGAAAGTACAGTCTTGGGTGTGGTCAGTCCAGAC	5160
Db	5120	CTTGGCAGCGCTGGGGTAAATTAGAGACAGAAAGTACAGTCTTGGGTGTGGTCAGTCCAGAC	5179
QY	5161	TGCCCCCAGGCAAGGCTTGTGGCTGTAGAAAAAGTCACGGCTTAGGCCCGGGCACGGTGGC	5220
Db	5180	TGCCCCCAGGCAAGGCTTGTGGCTGTAGAAAAAGTCACGGCTTAGGCCCGGGCACGGTGGC	5239
QY	5221	TCAGGCGCTATATCCACAGCATTTTGGGAGGCCGAGCGGGGTGTGATCACAGAGTCAAGAGA	5280
Db	5240	TCAGGCGCTATATCCACAGCATTTTGGGAGGCCGAGCGGGGTGTGATCACAGAGTCAAGAGA	5299
QY	5281	TCGTGACCATTCTGGCTAACAGGGTGAAGCCCGCTCTCTATTAATAAATACAAAAAATTTGG	5340
Db	5300	TCGTGACCATTCTGGCTAACAGGGTGAAGCCCGCTCTCTATTAATAAATACAAAAAATTTGG	5359
QY	5341	CCGGGCGATGTTGGCGGGGCACTGTATGTTCCAGCTACTCGGGAGGCTGAGCGAGAGATG	5400
Db	5360	CCGGGCGATGTTGGCGGGGCACTGTATGTTCCAGCTACTCGGGAGGCTGAGCGAGAGATG	5419
QY	5401	GGGTGAACCCCAAGAGGCAAGATTGTGCAGTGAGCCGAGATCGCGCACTGCATCCAGCT	5460
Db	5420	GGGTGAACCCCAAGAGGCAAGATTGTGCAGTGAGCCGAGATCGCGCACTGCATCCAGCT	5479
QY	5461	GGGGGACAGACAGATCCATCTCTGGAAAAAAGAAAAAGCTTCAGAGTGTGAGCCAGA	5520
Db	5480	GGGGGACAGACAGATCCATCTCTGGAAAAAAGAAAAAGCTTCAGAGTGTGAGCCAGA	5539

QY	5521	GGCCAGGCGTGAATTCGTCACTTACATACCTTGGGCGAAGGCACTCTTCCCTGGC	5580
Db	5540	GGCCAGGCGTGAATTCGTCACTTACATACCTTGGGCGAAGGCACTCTTCCCTGGC	5599
QY	5581	CCAGTTACCGGGGTGGATCGACTCCAGAGTCCCTTCCAGCATTTAAGCGTGCATGGTTC	5640
Db	5600	CCAATTCCAGGGGGTTGGATTCGATCCAGAGTCCCTTCCAGCATTTAAGCGTGCATGGTTC	5659
QY	5641	TAAATGAGAAAGATGGGGCAGTTTCCCTCTCTACCCCGACCCGCTGTCCACTTCAAGGT	5700
Db	5660	TAAATGAGAAAGATGGGGCAGTTTCCCTCTCTACCCCGACCCGCTGTCCACTTCAAGGT	5719
QY	5701	GAATGACAGGGAAGTACGTCGTCCCAATCCCGAGTTCCAAAGCCCTTGGGGACCCCTAC	5760
Db	5720	GAATGACAGGGAAGTACGTCGTCCCAATCCCGAGTTCCAAAGCCCTTGGGGACCCCTAC	5779
QY	5761	TGTCAGGCGTCGTGCACGAGAGGTGAAGGTCAGGTGAAGCCAAATGCGCTCGAAGGCTCTTG	5820
Db	5780	TGTCAGGCGTCGTGCACGAGAGGTGAAGGTCAGGTGAAGCCAAATGCGCTCGAAGGCTCTTG	5839
QY	5821	CCCTAATTCGGGACAGACATCCGGTTTCTCTGCGCTTACCGGGGATTTCTAAGGGGCTTTAGC	5880
Db	5840	CCCTAATTCGGGACAGACATCCGGTTTCTCTGCGCTTACCGGGGATTTCTAAGGGGCTTTAGC	5899
QY	5881	CGAATGAGATCATGCGGGGGCGGGGGGTTTCTGGGGGAGTTCCAGCTAATCAATTGGGA	5940
Db	5900	CGAATGAGATCATGCGGGGGCGGGGGGTTTCTGGGGGAGTTCCAGCTAATCAATTGGGA	5959
QY	5941	CAGGACAGCCTTGGAACCTTCGATGATGCTATTCACAAGTGTGGGGTGGGACAGACCCAA	6000
Db	5960	CAGGACAGCCTTGGAACCTTCGATGATGCTATTCACAAGTGTGGGGTGGGACAGACCCAA	6019
QY	6001	GACCCAAATGTCCTTATCTCAGGTAGGGGCTCAGAGGTCCTCCAGACAGCAGCCTCCGG	6060
Db	6020	GACCCAAATGTCCTTATCTCAGGTAGGGGCTCAGAGGTCCTCCAGACAGCAGCCTCCGG	6079
QY	6061	AGACTTTGGGGGTAGGAATGGGACCAACAGGCTTCTTTTCTCTTATGAATTTGGG	6120
Db	6080	AGACTTTGGGGGTAGGAATGGGACCAACAGGCTTCTTTTCTCTTATGAATTTGGG	6139
QY	6121	GCGTTGGGGGACAGGCTTGAGAAATCCAAAGGAGAGGGGCAAGAGACACTCCCCACAG	6180
Db	6140	GCGTTGGGGGACAGGCTTGAGAAATCCAAAGGAGAGGGGCAAGAGACACTCCCCACAG	6199
QY	6181	TCCTCCAGAGGAGAGAGAGGAGACCCCGACCTCAGCTGCACCTTCCCAAGAGGCT	6235
Db	6200	TCCTCCAGAGGAGAGAGAGGAGACCCCGACCTCAGCTGCACCTTCCCAAGAGGCT	6254
RESULT 3			
US-09-845-020A-6			
Sequence 6, Application US/09845020A			
Publication No. US20030022850A1			
GENERAL INFORMATION:			
APPLICANT: Treco, Douglas A.			
APPLICANT: Heartlein, Michael W.			
APPLICANT: Seiden, Richard F.			
TITLE OF INVENTION: Genomic Sequences for Protein Production			
TITLE OF INVENTION: and Delivery			
FILE REFERENCE: 50010/017003			
CURRENT APPLICATION NUMBER: US/09/845,020A			
CURRENT FILING DATE: 2001-04-27			
PRIOR APPLICATION NUMBER: US 09/305,384			
PRIOR FILING DATE: 1999-05-05			
PRIOR APPLICATION NUMBER: US 60/084,649			
PRIOR FILING DATE: 1998-05-07			
NUMBER OF SEQ ID NOS: 8			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 6			
LENGTH: 2834			
TYPE: DNA			
ORGANISM: Homo sapiens			

US-09-845-020A-6

Query Match 45.5%; Score 2834; DB 11; Length 2834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1890 CCGGACAGTCTCACAGCCCTGTTGCTCTGCGGGCCCTGCTGCTGGGCTCCGACATT 1949
DB 1 CCGGACAGTCTCACAGCCCTGTTGCTCTGCGGGCCCTGCTGCTGGGCTCCGACATT 60
QY 1950 GGTGGCACTTAGAGAGCCCTTACGCCACCGCTGACATGTGGAGACCCCTTCTGGGCTG 2009
DB 61 GGTGGCACTTAGAGAGCCCTTACGCCACCGCTGACATGTGGAGACCCCTTCTGGGCTG 120
QY 2010 GCCAAGGCGACAGCGGCTCCCTCAAGCTTGCAGGAGAGTGTGAGAGGAGAGGCTTAAGCA 2069
DB 121 GCCAAGGCGACAGCGGCTCCCTCAAGCTTGCAGGAGAGTGTGAGAGGAGAGGCTTAAGCA 180
QY 2070 GGAACCGGGGCTGGCGACAGGCGCTTGGCGGCACGTGAGATTCCGGTGGGCGTGGGCTT 2129
DB 181 GGAACCGGGGCTGGCGACAGGCGCTTGGCGGCACGTGAGATTCCGGTGGGCGTGGGCTT 240
QY 2130 GGGCGGCCCCGCACTCGGAGCAGCGGGCCAGCCCTGCGAGGCCCCGGCAATGAGAGGCT 2189
DB 241 GGGCGGCCCCGCACTCGGAGCAGCGGGCCAGCCCTGCGAGGCCCCGGCAATGAGAGGCT 300
QY 2190 TAGACCCGGGCGACAGCGGCTGGGGGTTACTGAGGTTGCCCCAGACAGTGGCAGCGCC 2249
DB 301 TAGACCCGGGCGACAGCGGCTGGGGGTTACTGAGGTTGCCCCAGACAGTGGCAGCGCC 360
QY 2250 GGGCGCTGTGCTGCTCGATTCTGACTGAGGCTTACAGCCCTTCCGCGGGGCGAGGGCTC 2309
DB 361 GGGCGCTGTGCTGCTCGATTCTGACTGAGGCTTACAGCCCTTCCGCGGGGCGAGGGCTC 420
QY 2310 GGGACCTGACAGCCCGCCATGCTTGAAGCTTCCCTTCATAGGCTCTGTTGGCGCCGAGCC 2369
DB 421 GGGACCTGACAGCCCGCCATGCTTGAAGCTTCCCTTCATAGGCTCTGTTGGCGCCGAGCC 480
QY 2370 TCCCGAGAGACACACCCCGCTGCTCCACAGCGGCACCTCCATTCGACAGCAGCAAGGCT 2429
DB 481 TCCCGAGAGACACACCCCGCTGCTCCACAGCGGCACCTCCATTCGACAGCAGCAAGGCT 540
QY 2430 GAGAAGTGGCGGCGCACGGCACCGGACTGGCAGGACGTAACCCCTGACGCCCTTGTCG 2489
DB 541 GAGAAGTGGCGGCGCACGGCACCGGACTGGCAGGACGTAACCCCTGACGCCCTTGTCG 600
QY 2490 GAATCCACTGGGTGAAGCCAGCTGGGCTCTGAGTCTGCTGAGAGACTTGAACCTTTA 2549
DB 601 GAATCCACTGGGTGAAGCCAGCTGGGCTCTGAGTCTGCTGAGAGACTTGAACCTTTA 660
QY 2550 TGTCTAGCTCAGGGATGTTAAATACCAATCAGACCCCTGTTCTAGCTAGAGGTCTGT 2609
DB 661 TGTCTAGCTCAGGGATGTTAAATACCAATCAGACCCCTGTTCTAGCTAGAGGTCTGT 720
QY 2610 GAATGACCAATCCACACTGTATCTAGTACTCTGATGGGCTTGGAGAACCTTTAT 2669
DB 721 GAATGACCAATCCACACTGTATCTAGTACTCTGATGGGCTTGGAGAACCTTTAT 780
QY 2670 GTCTAGCTCAGGATTTTAAATACCAATTCGACACTGTATCTAGCTCAAGGTTTGA 2729
DB 781 GTCTAGCTCAGGATTTTAAATACCAATTCGACACTGTATCTAGCTCAAGGTTTGA 840
QY 2730 AACAACCAATCAGACACCCCTGTGCTAGCTCAGGCTATGTGAATGACCAATTCAGAGTC 2789
DB 841 AACAACCAATCAGACACCCCTGTGCTAGCTCAGGCTATGTGAATGACCAATTCAGAGTC 900
QY 2790 TGTATCTGCTACTTTTCATGGGCTCCGTGTGAAGAGACCAAAAGGCTTTGTGGA 2849
DB 901 TGTATCTGCTACTTTTCATGGGCTCCGTGTGAAGAGACCAAAAGGCTTTGTGGA 960
QY 2850 GCAATAAAGCTTATACCTTGAGTGCAGGTGGGCTGAGTCCGAAAAGAGAGTCAAGGAA 2909
DB 961 GCAATAAAGCTTATACCTTGAGTGCAGGTGGGCTGAGTCCGAAAAGAGAGTCAAGGAA 1020

QY 2910 GGGAGATTAAGGTGGGCGCTTTTATAGATTTGGTAGTTAAAGAAAAATTACACTCA 2969
DB 1021 GGGAGATTAAGGTGGGCGCTTTTATAGATTTGGTAGTTAAAGAAAAATTACACTCA 1080
QY 2970 AGGGGTTTGTCTCTCGGGGCGCAGAGTGGGGGCTCGCAAGTGTCTAGTGGGGTGT 3029
DB 1081 AGGGGTTTGTCTCTCGGGGCGCAGAGTGGGGGCTCGCAAGTGTCTAGTGGGGTGT 1140
QY 3030 TTTTGAAGCAGAGATGAGCCAGGAAAAAGACTTTCACAAAGTATGTCATTAAGGCA 3089
DB 1141 TTTTGAAGCAGAGATGAGCCAGGAAAAAGACTTTCACAAAGTATGTCATTAAGGCA 1200
QY 3090 AGGACCCGCAATTTACACCTCTTTGTGTGAATGTCAATGATTAGTTGGGCGAGGC 3149
DB 1201 AGGACCCGCAATTTACACCTCTTTGTGTGAATGTCAATGATTAGTTGGGCGAGGC 1260
QY 3150 ATATTCACTTTCTTTGTGATTTCTACAGTACTTTCAGGCGCATGCGGCGTATATGTCAAG 3209
DB 1261 ATATTCACTTTCTTTGTGATTTCTACAGTACTTTCAGGCGCATGCGGCGTATATGTCAAG 1320
QY 3210 TTACAGGGGATGCGATGGCTTGGCTTGGCTCAGAGGCTTGAACGCTACTTGTGGGC 3269
DB 1321 TTACAGGGGATGCGATGGCTTGGCTTGGCTCAGAGGCTTGAACGCTACTTGTGGGC 1380
QY 3270 CTGGAGAAATGTTTGTGTGACACTGTATCTAGTTAATCTAGTGGGACGTGAGAAC 3329
DB 1381 CTGGAGAAATGTTTGTGTGACACTGTATCTAGTTAATCTAGTGGGACGTGAGAAC 1440
QY 3330 CTTTGTGCTACTCAGGGAATTGTAAGCAGCAATCAGCGCCGTGTCAAAACAGCCAC 3389
DB 1441 CTTTGTGCTACTCAGGGAATTGTAAGCAGCAATCAGCGCCGTGTCAAAACAGCCAC 1500
QY 3390 TCGGCTCTACCAATCAGCAGGATGTGGTGGGCGCAGATTAAGATAATAACAGGCTAC 3449
DB 1501 TCGGCTCTACCAATCAGCAGGATGTGGTGGGCGCAGATTAAGATAATAACAGGCTAC 1560
QY 3450 CCGAGCCAGCAATGGCAACGGCCACAGGTCCCTATTCACAAATATGGCAGCTTGTCTTT 3509
DB 1561 CCGAGCCAGCAATGGCAACGGCCACAGGTCCCTATTCACAAATATGGCAGCTTGTCTTT 1620
QY 3510 TGCTGTTTGGCAATATCTTGTACTGCTGCTGCTTTTGGGTGTCACACTCTTTTATGAGC 3569
DB 1621 TGCTGTTTGGCAATATCTTGTACTGCTGCTGCTTTTGGGTGTCACACTCTTTTATGAGC 1680
QY 3570 TGTAACTCTACACAGGAGTCTGCAGCTTCACTCTCTGAAAGCCACTAAGACCAAGAGCC 3629
DB 1681 TGTAACTCTACACAGGAGTCTGCAGCTTCACTCTCTGAAAGCCACTAAGACCAAGAGCC 1740
QY 3630 ACCGGAGGAATGAACAACCTCCGGCCGGCTGCTTTAAGAGTATTAACACTCACGGGAA 3689
DB 1741 ACCGGAGGAATGAACAACCTCCGGCCGGCTGCTTTAAGAGTATTAACACTCACGGGAA 1800
QY 3690 GGTGTGAGCTTCACTCTCAGCGCAGCGAGACCAAGAACCCCAAGAGGAAGAAATCC 3749
DB 1801 GGTGTGAGCTTCACTCTCAGCGCAGCGAGACCAAGAACCCCAAGAGGAAGAAATCC 1860
QY 3750 GAACACATCTGAACATCAGAAAGAACAACTCCAGATCAGCAGCACTTAAGAGCTGTAA 3809
DB 1861 GAACACATCTGAACATCAGAAAGAACAACTCCAGATCAGCAGCACTTAAGAGCTGTAA 1920
QY 3810 CTCACTGGAGGGTCCGGGCTTCTTGTGAAGTCACTGAGACCAACAGCACTCACAGTT 3869
DB 1921 CTCACTGGAGGGTCCGGGCTTCTTGTGAAGTCACTGAGACCAACAGCACTCACAGTT 1980
QY 3870 TCGGACACAAGCCAGGAGTTTGAGATCAGCTGGGCAACATGATGAATGCCCTCTGTG 3929
DB 1981 TCGGACACAAGCCAGGAGTTTGAGATCAGCTGGGCAACATGATGAATGCCCTCTGTG 2040
QY 3930 CAAAAAATAAATAATTCAAAAATTTGGCGGAGCATGGTGGCTGCTGGTGGCCAG 3989
DB 2041 CAAAAAATAAATAATTCAAAAATTTGGCGGAGCATGGTGGCTGCTGGTGGCCAG 2100

```

Oy 3990 CTACGGGAGGAGCTAAAGTGGAGAGATCGTTGAGCCCTGGGAGGTGAAGACTGCACTGAG 4049
    |||||||
Db 2101 CTACGGGAGGAGCTAAAGTGGAGAGATCGTTGAGCCCTGGGAGGTGAAGACTGCACTGAG 2160
    |||||||
Oy 4050 CTGTGATTTGACACAGCCCTCTAGGCTGGGGGAGACACTGAGACCCCTGTTTCCCTCCG 4109
    |||||||
Db 2161 CTGTGATTTGACACAGCCCTCTAGGCTGGGGGAGACACTGAGACCCCTGTTTCCCTCCG 2220
    |||||||
Oy 4110 CAAAAAATTTGACAAAAAGTAAATGAAGTGGTCCCTGATATGCTAGGCGCAGTGGCTCAT 4169
    |||||||
Db 2221 CAAAAAATTTGACAAAAAGTAAATGAAGTGGTCCCTGATATGCTAGGCGCAGTGGCTCAT 2280
    |||||||
Oy 4170 GCTGTAAATCCACACTTTTGGGAAGCCGAGGCGGCGGCTCACCTTAAGTCAAGAGTGT 4229
    |||||||
Db 2281 GCTGTAAATCCACACTTTTGGGAAGCCGAGGCGGCGGCTCACCTTAAGTCAAGAGTGT 2340
    |||||||
Oy 4230 GAGACCAAGCCCTGGCCAAACATGAGAAAGCCCATCTCTTCTAATAATACAAAAATTAGCCGG 4289
    |||||||
Db 2341 GAGACCAAGCCCTGGCCAAACATGAGAAAGCCCATCTCTTCTAATAATACAAAAATTAGCCGG 2400
    |||||||
Oy 4290 CTGTGGGGGAGAGTGGTGGAGCATGCTGTAAATCCACACTGCTAGGAGGCTGAGGAGCA 4349
    |||||||
Db 2401 CTGTGGGGGAGAGTGGTGGAGCATGCTGTAAATCCACACTGCTAGGAGGCTGAGGAGCA 2460
    |||||||
Oy 4350 GAATCACTTGAACCCAGAGGCGGCTGTGCACTGAGCGGAGATCGTCCATTTGCACTCC 4409
    |||||||
Db 2461 GAATCACTTGAACCCAGAGGCGGCTGTGCACTGAGCGGAGATCGTCCATTTGCACTCC 2520
    |||||||
Oy 4410 ACCCACTCCAGCCCTGGGCAACAGAGCCAACTGTGTCTTAAAAAATTAAGTGA 4469
    |||||||
Db 2521 ACCCACTCCAGCCCTGGGCAACAGAGCCAACTGTGTCTTAAAAAATTAAGTGA 2580
    |||||||
Oy 4470 CCTGACATATGAAGAGTGTGCAATGCAATAGTGGCAGGCAACATGTTAAGTGA 4529
    |||||||
Db 2581 CCTGACATATGAAGAGTGTGCAATGCAATAGTGGCAGGCAACATGTTAAGTGA 2640
    |||||||
Oy 4530 GCTCTGCTTCATATGCTCTGTAAATAAACCCACCTCAAGGCCAGGTGAGTGGCTCAT 4589
    |||||||
Db 2641 GCTCTGCTTCATATGCTCTGTAAATAAACCCACCTCAAGGCCAGGTGAGTGGCTCAT 2700
    |||||||
Oy 4590 GCTCTGCTTCATATGCTCTGTAAATAAACCCACCTCAAGGCCAGGTGAGTGGCTCAT 4649
    |||||||
Db 2701 GCTCTGCTTCATATGCTCTGTAAATAAACCCACCTCAAGGCCAGGTGAGTGGCTCAT 2760
    |||||||
Oy 4650 GAGACCAAGCCCTGAGCACCACCAATGTTGAATCCCACTCTAATAAATTAAGTGA 4709
    |||||||
Db 2761 GAGACCAAGCCCTGAGCACCACCAATGTTGAATCCCACTCTAATAAATTAAGTGA 2820
    |||||||
Oy 4710 TGAGCATGTGTGTG 4723
    |||||||
Db 2821 TGAGCATGTGTGTG 2834
    |||||||

```

```

RESULT 4
US-09-845-020A-7
: Sequence 7, Application US/09845020A
: Publication No. US20030022850A1
: GENERAL INFORMATION:
: APPLICANT: Treco, Douglas A.
: APPLICANT: Heartlein, Michel W.
: APPLICANT: Seiden, Richard F.
: TITLE OF INVENTION: Genomic Sequences for Protein Production
: FILE REFERENCE: 50010/017003
: CURRENT APPLICATION NUMBER: US/09/845,020A
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: US 09/305,384
: PRIOR FILING DATE: 1999-05-05
: PRIOR APPLICATION NUMBER: US 60/084,649
: PRIOR FILING DATE: 1998-05-07
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7

```

```

: LENGTH: 1252
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-845-020A-7

Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

20.1%; Score 1252; DB 11; Length 1252;

Oy 4728 CCTGTAATCCACACTTACTTGGAGGCTGAGGCGAGGAAATCACTAGAACCAAGGAGCGG 4787
    |||||||
Db 1 CCTGTAATCCACACTTACTTGGAGGCTGAGGCGAGGAAATCACTAGAACCAAGGAGCGG 60
    |||||||
Oy 4788 AGGTGTGTGAGACCGGAGATGCTGCACTTTCACACTCCAGCTGAGCAATGAGCGAAATCC 4847
    |||||||
Db 61 AGGTGTGTGAGACCGGAGATGCTGCACTTTCACACTCCAGCTGAGCAATGAGCGAAATCC 120
    |||||||
Oy 4848 ATCTCAAAAAACAACAACAAACCACTCTCTACTCCAGGAGCTGGGTACAGACT 4907
    |||||||
Db 121 ATCTCAAAAAACAACAACCAACCACTCTCTACTCCAGGAGCTGGGTACAGACT 180
    |||||||
Oy 4908 GGGCCACATCAAGTGTGCTGAGCCACAGAGCTAAGGCGGAGCTGACAGACCGGGA 4967
    |||||||
Db 181 GGGCCACATCAAGTGTGCTGAGCCACAGAGCTAAGGCGGAGCTGACAGACCGGGA 240
    |||||||
Oy 4968 CCAGATTAACAGTGTGAGATCACTGTGTGAGATGAGACGTCCCTGCTGATGATGACAC 5027
    |||||||
Db 241 CCAGATTAACAGTGTGAGATCACTGTGTGAGATGAGACGTCCCTGCTGATGATGACAC 300
    |||||||
Oy 5028 CAGGGGGCCCCCAACACACAGAGATGGCCCATCCAGTCCACATCCACTTCTCATCA 5087
    |||||||
Db 301 CAGGGGGCCCCCAACACACAGAGATGGCCCATCCAGTCCACATCCACTTCTCATCA 360
    |||||||
Oy 5088 GAGATGTCTGTTCTTGGCAGCGCTGGGTTAAATTAAGAAGAGTGAAGTCTTGGGTG 5147
    |||||||
Db 361 GAGATGTCTGTTCTTGGCAGCGCTGGGTTAAATTAAGAAGAGTGAAGTCTTGGGTG 420
    |||||||
Oy 5148 TGGTACATCAGACGCTCCGAGGAGGCGCTGTTGGCTGTAGAAAAGTTCAGGCTTAGGC 5207
    |||||||
Db 421 TGGTACATCAGACGCTCCGAGGAGGCGCTTGTGGCTGTAGAAAAGTTCAGGCTTAGGC 480
    |||||||
Oy 5208 CGGGCAGGCTGCTCAGCCTGTAAATCCACACTTTTGGAGGCGGAGCGGCTGATCA 5267
    |||||||
Db 481 CGGGCAGGCTGCTCAGCCTGTAAATCCACACTTTTGGAGGCGGAGCGGCTGATCA 540
    |||||||
Oy 5268 CGAGGTCAAGAGATGTGACCATCTGCTTAACACAGTGAACCCCGTCTACTAAAAA 5327
    |||||||
Db 541 CGAGGTCAAGAGATGTGACCATCTGCTTAACACAGTGAACCCCGTCTACTAAAAA 600
    |||||||
Oy 5328 TACAAAAAATTTGGGCGGAGATGTGGGCGGACCTGTAGTTCCAGCTACTGCGGAGGCTG 5387
    |||||||
Db 601 TACAAAAAATTTGGGCGGAGATGTGGGCGGACCTGTAGTTCCAGCTACTGCGGAGGCTG 660
    |||||||
Oy 5388 AGCGAGAGAAATGCGGTGAACCCGAGAGGAGAGTTCAGTGAAGCCGAGATCGCGCAC 5447
    |||||||
Db 661 AGCGAGAGAAATGCGGTGAACCCGAGAGGAGAGTTCAGTGAAGCCGAGATCGCGCAC 720
    |||||||
Oy 5448 TGCACCTCAAGCTGGGCGACAGAGCAAGACTCTGAGAAAAGAAAAGAAAGCTTCA 5507
    |||||||
Db 721 TGCACCTCAAGCTGGGCGACAGAGCAAGACTCTGAGAAAAGAAAAGAAAGCTTCA 780
    |||||||
Oy 5508 GGTGTGAGCCAGAGGCGGAGGCTGTAAATTTGTCCTTACATGACCTTGGGCAAGGAGC 5567
    |||||||
Db 781 GGTGTGAGCCAGAGGCGGAGGCTGTAAATTTGTCCTTACATGACCTTGGGCAAGGAGC 840
    |||||||
Oy 5568 TTCTCTTCCCTGGCCAGTTCAAGCGGGGTGGAATGCACTCAAGAGTCCCTTCAGATTTA 5627
    |||||||
Db 841 TTCTCTTCCCTGGCCAGTTCAAGCGGGGTGGAATGCACTCAAGAGTCCCTTCAGATTTA 900
    |||||||
Oy 5628 CGCTGCAATGTTCTTAAGATGAAGAGTGGGCAAGTTTCCCTCTTCAACCCAGCCGCTG 5687
    |||||||
Db 901 CGCTGCAATGTTCTTAAGATGAAGAGTGGGCAAGTTTCCCTCTTCAACCCAGCCGCTG 960
    |||||||

```


Db	2865	TTGTGCAAAACTCTGTATATCTAATTAATCTGATGGGAGCTTGGAGAACCTTTGTATTCTAAC	2806
QY	3343	TCAGGAGATTGTAAGACGACCAATGACGCCCTGTCCAAMACAGACCACTCGGCTCTACCAA	3402
Db	2805	TCAGGAGATTGTAAGACGACCAAGTCAGCGCCCTTGACAAAACAGGGCCGCTTACCAA	2746
QY	3403	TCAGCAGATGTGGGTGGGGCCAGATTAAGAAATTAAGACAGGCTGCCCGACAGAGT	3462
Db	2745	TCAGCAGGATGTGGGTGGGGCCCAATTAAGAAATTAAGACAGGCTGCCCGACAGATTT	2686
QY	3463	GGCAACGGCAGAGTCCCTATTCACAAATATGGAGAGCTTTGTCTTTTGGCTGTTTGGAT	3522
Db	2685	GGCAATCAGCTCGGGTCCCTTGCCATGACACAGTGGAGGCTTTGTCTTTCGCTTTTGCAT	2626
QY	3533	AAATCTTGCTACTGCTGCTTTTGGGTCCACACTGCTTTTATGAGCTGTAACTACCTACC	3582
Db	2625	AAATCTTGCTACTGCTGCTTTGGGTCCACGCTGCTTTTGTAACTATTAACACTACT	2566
QY	3583	ACGAAGGCTGTCAGCTTACCTGTCGAAAGGCCACTTAAGACACAGACCCACCGGGAGGAATG	3642
Db	2565	GCAAGATCTCAGCTTACCTGCTGAGGCCAGGACGACACGACCCCACTGGGAGGAACG	2506
QY	3643	AACAACCTCGGCGCGCTGCTCTTAAGAGCTATTAAGACTCTACCGGGAAGGTCTGCAGCTTC	3702
Db	2505	AACAACCTCCAGATCGCGCTGCTCTTAAGAGCTGTAACTACTCACTGGGAAGGTCTGCAATTC	2446
QY	3703	ACTGCTCAG -CCACGAGAACCCAGACACCCACCAAGAGAAAGAAATCGGAACACATCTGA	3761
Db	2445	ATTCCTGAGCCACGAGACGACCAACCCACCAAGAGAAAGAAATCGGAACACATCTGA	2386
QY	3762	ACATCAGAGAGAACAACTCCAGATGACACCACTTAAGAGCTGTAACTACTCTGCGAGG	3821
Db	2385	ACATCAGAGAGGACAGACTCAAGACGTCGACCTTAAGAGCTGTAACTACTCTGCGAGG	2326
QY	3822	GTCGCGGCTTCCTTTTGAAGTCAATGACCAAGCACTCACCAGCTTGCGACACAAGC	3881
Db	2325	GTCGCGGCTTCATTTCTTGAAGTCAATGACCAAGCAAGCAACCCCAATTCGAGACATATTA	2266
QY	3882	CCAGAGCTTTGAG 3894	
Db	2265	CTATTACTAATAG 2253	
RESULT 6			
US-10-025-966A-23/C			
: Sequence 23, Application US/10025966A			
: Publication No. US20030148920A1			
GENERAL INFORMATION:			
APPLICANT: Rosen, Steven D.			
APPLICANT: Palmeri, Diana			
APPLICANT: Stefan, Hemmerich			
TITLE OF INVENTION: Sulfatases and methods of use			
TITLE OF INVENTION: thereof			
FILE REFERENCE: UCA1230			
CURRENT APPLICATION NUMBER: US/10/025, 966A			
CURRENT FILING DATE: 2002-12-21			
PRIOR APPLICATION NUMBER: 60/258, 577			
PRIOR FILING DATE: 2000-12-27			
PRIOR APPLICATION NUMBER: 60/267, 831			
PRIOR FILING DATE: 2001-09-02			
NUMBER OF SEQ ID NOS: 26			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 23			
LENGTH: 260209			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (1)...(260209)			
OTHER INFORMATION: n = A,T,C or G			
US-10-025-966A-23			

Best Local Similarity 69.6%; Pred. No. 8.5e-186;
Matches 1399; Conservative 0; Mismatches 243; Indels 369; Gaps 14;

QY	1875	ACAGGTGACAGCATGCGGGCAGTCCCTCACAGCCGCTGTTGGCTGCGGCGCCCTCTGCG	1934
Db	222481	ACTCGTAGACGCTGCTGGCACTCTCTCACAGCCCTGCTGCTCTCCGGCCTCTCTGC	2224222
QY	1935	CTGGGCTCCCACTTGGGTGGCACTTGAAGAGCCCTTCA--GCCACCGCTGCACCTGGGA	1993
Db	222421	GTAGGCTCCCACTTGGCGCGCACTTGAAGAGCCCTTCAAGGCCCTTCAAGGCCCTGACCTGGGA	2223622
QY	1994	GCCCCCTTTCCTGGGCTGAGCCCAAGGCCAGACCCGGCTCTCCCTCACGTTTCAGGAGGTGGGA	2053
Db	222361	GCCCTTTCCTGAGCTGGCCAAAGGCCAGTCCCGGCTCTCCCTCACGTTTCAGGAGGTGGGA	2223020
QY	2054	GGAAGAGGCTCAACGACGAACCGGGGCTCGGACAGGCGGCTTGGCGGGCCAGCTGGATTCC	2113
Db	222301	GGGAGAGGCGCAATGGGAACCGGGGCTCTGTCGGTGTCTTGGCGGGCCAGCTGGATTCC	2224242
QY	2114	GGGTGGCGGTGGCTTGGCGGGCCCGCACTCGAGACAGCGGCGCAAGCCCTGCCAGGCC	2173
Db	222421	GGATGGGCGCTGGCTTGGCTGGCCGCCACACTCGGACAGAGCCCGGCGCCCTTGGC--GGCC	2221830
QY	2174	CGGGCAATAGAGGCTTACACCCCGGGCCAGGGGCTGGGGAGGGGTACTGGGTGCCCA	2233
Db	222182	CGGGCAGTAGAGGGGCTTACACCCCGGGCCAGGGGCTGGGGAGGGGTACTGGGTGCCCA	2221350
QY	2234	GCAGTGCACCGCCGCGCGGCTGTGTGCTCGCTGATTTCACACTGAGGGCTTACAGGCTTC	2293
Db	222134	GCAGTAGCACGGCCACCGGCGCTG---GCCTGATTTCACCGGGCCTTAGCTGCTTC	222079
QY	2294	CCGCGGGGACGGGCTCGGAGCTGCAGCCCGCATGCTGAGCTCC-----CTCCATG	2348
Db	222078	CCGACAGGGGACGCTCTCGGGA--CTGCAGCCCGCATGCTCGAGGCTCCACCCACTCATG	222020
QY	2349	GGCTCGTGGCGGCCGAGCCTGCCGAGAGACACCCCGTGCACAGAGGCCAGT	2408
Db	222019	GGCTCTGTGACGCGCCGAGCCTGCCGAGAGACACCCCGTGTCTATGGCTCCCAT	221960
QY	2409	CCCATGCACGCAAGGGCTGAGAAGTGCGGGCGACGCGACCGGACTGCGAGGACG	2468
Db	221959	CCCATGCACACCCAAAGGCTGAGAGTCTGAGCGCATGGC--GCGGAGCTGCGACAGCGC	2219010
QY	2469	TACCCCTGAGGCCCTGTGCCGAACTCCATGGGTGAAGCAGCTGGGCTTCGATGCTGG	2528
Db	221900	TCACCTGTAGGCCCTGTGCCGAACTCCATGAAGCAGCTGGGCTTCGATGCTGG	2218411
QY	2529	TGGAAGCTTGGAGAACTTATGTGTAGCTCAGAGATCTTAATATACCAATCAGACCC	2588
Db	221840	TGGGAGCGTGGAG--TCTTATGTCTAGCTCAGGGATTGTAAACACCAATCAGATCC	2217830
QY	2589	TGTGTCTAGCTCAGGGTCTGTGAATGCACCAATCCACACTGTGTATCTAGCTCTGAT	2648
Db	221782	TGTGTCTAGCTCAGGGTGTGTGAAGTGTAGTGCACCAATGTGACACTGTGTATGTAGCTCTGAT	2217230
QY	2649	GGGGGCTTGGAGAACTTATGTGTAGCTCAGGAGTGTAAATATACCAATCGGCACTCT	2708
Db	221722	GGGGGCTTGGAGAACTTATGTGTAGCTCAGGGA-----	2216888
QY	2709	GTATCTAGCTCAAGTTGTGAATACACCAATCAGACCCCTGTGTAGCTCAGGGTATG	2768
Db	221687	-----TTGTAAACACCAATCAGACCCCTGTGTATGTGAATAGGTTTG	2216444
QY	2769	TGAATGCACCAATCAGACGCTGTATCTGTGCTACTTCAATGGGCATCCGTGTCAAGAGAC	2828
Db	221643	TGACTGCACCAATCAGACA--CTGTATCTGTGCTGTCTGTGGG-----	2216030
QY	2829	CACCAACAAGGCTTGTGTGAGCAMPAAAGCTTATATACCTGGGTGAGGTGGGTGAG	2888
Db	221602	-----	2216030
QY	2889	TCGGAAGAGAGTACGGMAAGGAGATTAAGGTGGCGCGCTTATTAAGGATTGGGTAG	2948


```

; NAME/KEY: Intron
; LOCATION: (37600)...(58973)
; OTHER INFORMATION: Intron 4
; NAME/KEY: Intron
; LOCATION: (59115)...(61405)
; OTHER INFORMATION: Intron 5
; NAME/KEY: Intron
; LOCATION: (61540)...(63027)
; OTHER INFORMATION: Intron 6
; NAME/KEY: Intron
; LOCATION: (63127)...(63463)
; OTHER INFORMATION: Intron 7
; NAME/KEY: Intron
; LOCATION: (63564)...(67187)
; OTHER INFORMATION: Intron 8
; NAME/KEY: Intron
; LOCATION: (67335)...(87922)
; OTHER INFORMATION: Intron 9
; NAME/KEY: Intron
; LOCATION: (88037)...(89288)
; OTHER INFORMATION: Intron 10
; US-10-002-491-10
```

Query Match 11.0%; Score 683.2; DB 14; Length 91000;

Best Local Similarity 89.2%; Pred. No. 1.1e-169; Mismatches 83; Indels 13; Gaps 5;

```

; Matches 796; Conservative 0; Mismatches 83; Indels 13; Gaps 5;
QY 1873 TGAAGGTGACAGAGTCCGCGAGTCCCTACAGCCCTCGTTCGCTCGCGCCCTCTCT
DB 85237 TGAGAGGTGACAGAGTCCGCGAGTCCCTACAGCCCTCGTTCGCTCGCGCCCTCTCT
QY 1933 GCCTGGGCTCCCACTTGGTGGGCACTTGAGAGAGCCCTTACGCCCTCAGCTGTGGG
DB 85297 GCTGGGCTCCCACTTGGTGGGCACTTGAGAGAGCCCTTACGCCCTCAGCTGTGGG
QY 1993 AGCCCTTTCTGGGCTGGCCAGGCAAGGCGCTCCCTCAGCTTGAGAGAGGTGGG
DB 85357 AGCCCTTTCTGGGCTGGCCAGGCAAGGCGCTCCCTCAGCTTGAGAGAGGTGGG
QY 2053 AGGAGAGGCTCAAGAGAAACGGGGCTGGCAGCGGCTTGGGGGCACTGAGATTG
DB 85417 AGGAGAGAGCGGAGCGGAAACGGGGCTGGCAGCGGCTTGGAGGCACTGGAATTG
QY 2113 CGGGTGGGCTGGGCTGGGCGGCGCCGCACTCGGAGAGAGGGGCGCAGGCGC
DB 85477 CGGGTGGGCTGGGCTGGGCGGCGCCGCACTCGGAGAGAGGGGCGCAGGCGC
QY 2173 CCGGGCAATGAGAGGCTTACACCCGCGGAGCGGCTGCGGAGAGGTGACTGGGTCGCC
DB 85536 CCGGGTAAATAAGGACTTAGACCCGCGGAGCGGCTGCGGAGAGGTGACTGGGTCGCC
QY 2233 AGCAGTCCAGCCCGCGGCGTGTCTCGCTGATTTCTGACTGGGCTTACGAGCCTT
DB 85596 AGCAGTCCAGCCCGCGGCGTGTCTCGCTGATTTCTGACTGGGCTTACGAGCCTT
QY 2293 CCGGCGGGGAGGCGGCTGGGAGCTGACAGCCCGGAGCGGCTGAGGCTGCC-----CTCAT
DB 85652 CCGGAGGGGAGGCGGCTGGGAGCTGACAGCCCGGAGCGGCTGAGGCTGCCAT
QY 2348 GGGCTCTGTGCGGCGCGAGCTCCCGAGAGCAGCACCCTCGCTCCACAGAGGCCAG
DB 85712 GGGCTCTGTGCGGCGCGAGCTCCCGAGAGCAGCACCCTCGCTCCACAGAGGCCAG
QY 2408 TCCCATGACAGCAGCAGGCTGAGAAAGTGGGCGGAGCAGGCAAGGAGTGGCAGGAGC
DB 85772 TCCCATGACAGCAGCAGGCTGAGAAAGTGGGCGGAGCAGGAGC -GCATGACTGGCAGAGC
QY 2468 CTACCCCTGACAGCCCTGTGCTGGGAGTCACTGGGTGAGCAGGCTGGGCTCTGAGTCTG
DB 85831 CTCACCTGACAGCCCTGTGCTGGGAGTCACTGGGTGAGCAGGCTGGGCTCTGAGTCTG
QY 2528 GTGGAGACTGGAGAACTTATGTCTAGTCAAGGAGTCTGAAATACCAATACAGCACC
; 2587
```

```

DB 85891 GTGGGAGCTGAGAGTCTTTATATCTAGCTCAGGAGTAAACACCAATACAGCACC
QY 2588 CTGTGTAGCTCAGGAGTCTGTGATGACCAATCAGCAGTGTATGACTAGCTGTA
DB 85951 CTGTGTAGCTCAGGAGTGTGTGAGTGCACCAATGACA--CTGTATGAGTGTCTG
QY 2648 TGGGCTTGGAGAACCTTATGTCTAGCTCAGGAGTGTAAATACCAATGCGCAGTC
DB 86009 TGGGCTTGGAGAACCTTATGTCTAGCTCAGGAGTGTAAATACCAATGAGCAGCC
QY 2708 TGTATCTAGCTCAGGAGTGTAAACACCAATACAGCAGCAGCTGAGTCT
DB 86069 TGTGTAGCTCAGGAGTGTGTGAAATGACCAATGACAGCAGCTGATCTAGCT
; 86120
```

RESULT 9

```

; US-10-012-984-14
; Sequence 14, Application US/10012984
; Publication No. US20030118561A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPID SCRAMBLASE 4 EXPRESSION
; FILE REFERENCE: RTS-0334
; CURRENT FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 92
; SEQ ID NO 14
; LENGTH: 104000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14992
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14993
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14994
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14995
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14996
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14997
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 14998
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15000
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15001
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15002
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15003
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15004
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 15005
; OTHER INFORMATION: unknown
```

NAME/KEY: unsure
 LOCATION: 15006
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15007
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15008
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15009
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15010
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15011
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15012
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15013
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15014
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15015
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15016
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15017
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15018
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15019
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15020
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15021
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15022
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15023
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15024
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15025
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15026
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15027
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15028
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15029
 OTHER INFORMATION: unknown
 NAME/KEY: unsure

LOCATION: 15030
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15031
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15032
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15033
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15034
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15035
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15036
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15037
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15038
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15039
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15040
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15041
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15042
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15043
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15044
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15045
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15046
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15047
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15048
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15049
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15050
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15051
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15052
 OTHER INFORMATION: unknown
 NAME/KEY: unsure
 LOCATION: 15053
 OTHER INFORMATION: unknown

Query Match

10.7%; Score 670; DB 14; Length 104000;


```

; APPLICANT: Brenner, Michael B
; APPLICANT: Behar, Samuel M
; TITLE OF INVENTION: Soluble CD1 Compositions and Uses Thereof
; FILE REFERENCE: B00801/70212
; CURRENT APPLICATION NUMBER: US/09/874,470
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,416
; PRIOR FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 5
; LENGTH: 10351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-874-470-5

Query Match      10.5%; Score 652.8; DB 9; Length 10351;
Best Local Similarity 58.4%; Pred. No. 3.4e-162;
Matches 1626; Conservative 0; Mismatches 942; Indels 218; Gaps 19;

QY 73 ATTATGCCGGGCGATGTTGGCTCAGCTCTGTAATCCCTGAACTTTGGACATCAAGCAA 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2709 AATCAGGCCAGGCGAGGTGGTCACACCCGTATCCAGCACTTTGGGAAGCCAAAGGCGAG 2650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 133 GTGATCACTTGAGTCAAGAGTTGAGACTGACCTGGCCAACTGGTAAGAACCTATCT 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2649 GGGAA-----TTGAGAGCCAGCCCTTGCCAACTGGCGAAACCCCTGTCT 2607
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 CCACATAAAATATACAAAATTTAGCCAGGCGATGTTGGACGACCTGTATCCCGGCTACT 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2606 CTACCAAAAAAAGTAAATTTAGCCAGGCGCATGATGTGCTGTATTTCCAGCTACT 2547
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 CAGAGGCTGAGGCGAGAGATCACTTGAACCCAGAGGCGGAGGTTGCATGAGCTGAG 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2546 CAGGAGGCTGAGGCGAGAGATTTGCTTGAGCCTAGAGAGTGGAGATTTGAGGAACTGAG 2487
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 ATCAACACCACTGCATCCAGCCTGGGTGACAGAGCAAGCTCTATCTCAAAAAATATAA 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2486 ATTGTGCCACTGTACTACATCTCTGGGCGAGAGAGTGTACTGTGTTCAAAAAAATAA 2427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 AAAATAAAAAATTTAGCCAGGCGATGTTAGTCACACCTTACTAGCTACTCAGGAGG 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2426 AAGAGGGAATTTTAAATTCATTGCAAAAGTAA----- 2392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 CTGAGGTGGAGGATCACTTGAACCTGGGGGAGTCAAGGCTACAGTGAAGCAAGATCATG 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2391 ----- 2392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 CCACATACACTCAGGCTGGGCAACAGAGAGACCCCTGTCTAAAAAATAATATAT 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2391 -----ATTAAGCAAAAGTATATTTAGAGAAATATAAACAATGTTGTAT 2350
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 553 AAAGAAAAAAGAGCTGTATATGTCCTGTGTCATACATACATACATATATATAT 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2349 CCACAAAAAATTTGACCTTTTATGTAAGTAAGATAAACACACCTAGTTTAAATTTTGA 2290
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 613 TGCAAATCTCAAGATCAAGATAGTCAATTTTAAAGCTTGTGGCCGTGATGCTCTCTGTC 672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2289 GGCCTACAGAACTCAAGATTAAGTCAAAATTTACTGACAAAGTAAATTAATCTGACAG 2230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 673 ACAATACACTGCGCCCTGTCTTCTAGCACAAAGACAGCATATAACATACATACATGAT 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2229 CAACCCCATATCATTAACAAAATATATGAAATTCCTTAATTTAAATTAATCCAGAA-- 2172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 733 TTTTATAGACATCGAGATTTGAAATTCATATGATTTTACATTTTAAATTAATCT 792
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2171 ----GATAAATTTCCAAAGAAAAATATATATTTAATTTTAAATTTTGAATATATGT 2116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 793 TTAATAATTTTCCCTTAACATTTAAAGTATAAAGCCGCGCAGCGCCATCTGTCAG 852
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2115 ACAATTAATTTCTTCTGCTTTACAAAAATTCATCATAGTGGCGCAGGACAGTGGCTCACT 2056
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 853 CCTGTAATTCAGACACTTTGGAGGCTGAGTGGGCAAGTCACTTGAGATCAACAGTTGCG 912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 2055 CCTGTAATCCAGCACTTTGGAGGCGGAGGAGGAGATCA--GGAGTCAAGGATGCG 1998
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 913 AGACGACCTGGGCCAATAGCAAAACCCC-ATTTCTATAAATAAATAAATTAATGCTG 971
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1997 AGACCATCTGCTGTAACAGGTGAACCCCTCTACTATAAATACAAAAAATTAAGCAG 1938
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 972 GGCATATGTCACACCTGTATCCACTACTTTGGGAGGCTGAGGCGAGGAATGCGT 1031
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1937 GCGGTGTGGCAGGTGCTATGTGCTCCACTATTCAGAACTGAGGCAAGAAATGCGT 1878
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1032 TGAACCTGGGAAGCGGAGGTTGACAGGACCAACATCATCCACTGACATCCAGCTGGG 1091
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1877 TGAACCTGGGAGGAGAGGTTGACAGGACCAAGATGTGCCCAGCTGACATCCAGCTGGG 1818
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1092 TGAACAGTGAAGACTTGTCTCAACGAAAAAAGTAAAGCCATTCATATTCAG 1151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1817 CAACAGAGAAAGACTCCGTCTCAAAAAA-----AAAAATATTCATCAGTGC 1762
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1152 TGTACATCACTGATCACTACAGTGTGCTGCTCTGAGGACATACCTGAGAACT 1211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1761 CTAAACCAATTTTAACTAAATAATCTATGCTGTGTGGCCAGAGGCTCTTTAAAG 1702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1212 AGAGTGTGCTGGTCAAGAGACATACATTTCCACATTAAGTAAAGACTACCAAGTTGCG 1271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1701 AATTTCCCTCTTATGTTTCTGAGGCTGCCCGCATGACATGTGCATCCCTTTG 1642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1272 ATCCAGAGAGGTTTTTTTTTACAAATCTACACTCCCCCGACGAAACAAATAGAGTTATC 1331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1641 CCAAAACCAAAATTTCTCAAAACATTTTAAAGATGTC-----CTTTGACAAAC 1589
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1332 CAGATCTTTACAAAGATCTTAAGCCAGTACAGATGAAAGAAACGGAAGTGGAGGG 1391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1588 TTAATCTTTCTATGTAAGATTTATGAGATTAACATTAATTAATGATTTAAATMAA 1529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1392 AAGCTGCCAGCCCTCTCAACCATGAAGAAATACCTGTAAGACCTTGAGATGCTGGAA 1451
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1528 ATTTTGTAGTACATTTGTCAAAATGTAAGAAATGTTCCATCTGCTAATTTTTTTCAAA 1469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1452 GGTAAATTAACGGGGGCTCTGTGAGGCTGCCCTGTGAGTACATCTGACTTCGAGCC 1511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1468 GAATTAATTAAGAGATTTATTTATTTGCTTATTTTCAAGAAAGATTCGAATTTCTACA 1409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1512 TCCAGTCCAGTCCAGCCCATGTGTCAATGAGCCAGTGAATGAGCCCTCAGCTCTGTT 1571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1408 TACTTAATTTCTGTTTGTATATCTTACTTCAATTTCTTAACATTTAATTAATTTT 1349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1572 TGG---TCTTTATTCCTCCCATGTGGGCTGAAGTCTGATGAGCGCTTATTCAGATG 1628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1348 CAGCTTTAGTAGAGTACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1629 TACAGCTTCTTGACAGAAAGTATGTACAGAAACACAGAGGCTTGGCAAGATATC 1688
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1288 CTGAGGCTGTAATTTAAATATGTAAGTAATTAATTAATTAATTAATTAATTAATTA 1229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1689 TAATGCAAAATCTTACCTGGCTGAGCCAGCAGTATCTGTGATCTTGAACAAGTTTT 1748
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1228 TTTTATTAATTTGGTATATTAAGTGCACAAACATGATATATATATATATATATAC 1169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1749 TCACCTCTGAGGCGCATCCCTGGCTACAAACACACAGTGGTGAAGAGTAATGA 1808
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1168 TGTGAAGTAAATGCCACAGTCAAGTATTAATCAATATCACTTACATAGTATGATTT 1109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1809 CGAAGTCCCTTACACCTGTATATCCAGCACTTTGGAGGCGCAAGCGGGTGA--TGGCT 1867
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1108 TTTCTTTTCTTTTCTTTTGAAGAGATTCACATCTGACCCAGGCTGGAGTGCAGT 1049
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1868 GAGCTGAGAGGTGAGACAGATGCGGCACTCTCAGAGCCCTGCTGCTGCGGCGCT 1927
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1048 GGCATTTGAAGGTGAGAGGCTGCTGCG--GTCCTCAGAGCACTTGTGCTGCGGGCT 990
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1928 CCTGTGCCGGGCTCCCACTTGCGTGGGCACTTGAAGAGCCCTTCAAGCCACGCTGCACT 1987
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 989 CCTGTGCTGGGCTCCACTTTTGGCAGCACTTGAAGAACTTTCAGCCCGCGCTGCACCT 930
QY 1988 GTGGAGAGCCCTTCTTGGGCTGGCCAGAGCCAGAGCCGCTCCCTCAGTTGACGAGG 2047
Db 929 GTAGAGCCCTTCTTGGGCTGGCCAGAGCCGAGCCGCTCCCTCAGTTGACGAGG 870
QY 2048 TGTGGAGGAGAGGCTCAGCAGAGAACCGGCGCTGGCAGCGCGCTTGGCGCCAGCTGG 2107
Db 869 TGTGGAGGAGAGGCGGAGCAGCACTGGCGCTGGCGGCGG-----GTTCCAGCTGG 818
QY 2108 AGTCCGAGTGGGCTGGGCTGGGCGCGCCGCACTCGGAGCAGCGGCGCAGCCCGCC 2167
Db 817 AGTCCGAGTGGGCTGGGCTGGGCGCGCCGCACTCGGAGCAGCGGCGCAGCCCGCC 759
QY 2168 AGGCCCCGGGCAATGAGAGGCTTACAGCCCGGCGCAGCGGCTGGAGGAGTGTACTGGT 2227
Db 758 -GGCCCCGAGCAATGAGAGCTTACAGCCCGGCGCAGCGGCTGGAGGAGTGTACTGGT 700
QY 2228 GCGCCAGCACTGGCAGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2287
Db 699 CCGCCAGCACTGGCAGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 2288 GCGCTCCCGGCGGCGGAGGCTCGGAGCTCGAGCGCGCAGCTGAGCTGCTGAGCTGCTG 2342
Db 643 GCGCTCCCGGCGGCGGAGGCTCGGAGCTCGAGCGCGCAGCTGAGCTGAGCTGAGCTGAG 584
QY 2343 TCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2402
Db 583 TCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 524
QY 2403 CCGAGTCCCATGCAACCAAGGAGGCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 2462
Db 523 CCGAGTCCCATGCAACCAAGGAGGCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 465
QY 2463 GGCAGCTACCCCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2522
Db 464 GGCAGCTACCCCTGAGCCCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 405
QY 2523 GTCGTGGAGACTTGGAGAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2572
Db 404 GTCGTGGAGACTTGGAGAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
QY 2573 -----ACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2601
Db 344 GCATTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
QY 2602 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2661
Db 284 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225
QY 2662 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2712
Db 224 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 165
QY 2713 CTAGCTCAAGGCTTGTAAACACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTG 2772
Db 164 CTAGCTCAAGGCTTGTAAACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTG 108
QY 2773 TGCACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2798
Db 107 TGCACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 82

RESULT 11
US-10-081-327-40
; Sequence 40, Application US/10081327
; Publication No. US20030129602A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UO01523
; CURRENT APPLICATION NUMBER: US/10/081,327
; CURRENT FILING DATE: 2002-02-21

QY 1048 TGTGGAGGAGAGGCTCAGCAGAGAACCGGCGCTGGCAGCGCGCTTGGCGCCAGCTGG 2107
Db 869 TGTGGAGGAGAGGCGGAGCAGCACTGGCGCTGGCGGCGG-----GTTCCAGCTGG 818
QY 2108 AGTCCGAGTGGGCTGGGCTGGGCGCGCCGCACTCGGAGCAGCGGCGCAGCCCGCC 2167
Db 817 AGTCCGAGTGGGCTGGGCTGGGCGCGCCGCACTCGGAGCAGCGGCGCAGCCCGCC 759
QY 2168 AGGCCCCGGGCAATGAGAGGCTTACAGCCCGGCGCAGCGGCTGGAGGAGTGTACTGGT 2227
Db 758 -GGCCCCGAGCAATGAGAGCTTACAGCCCGGCGCAGCGGCTGGAGGAGTGTACTGGT 700
QY 2228 GCGCCAGCACTGGCAGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2287
Db 699 CCGCCAGCACTGGCAGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 2288 GCGCTCCCGGCGGCGGAGGCTCGGAGCTCGAGCGCGCAGCTGAGCTGCTGAGCTGCTG 2342
Db 643 GCGCTCCCGGCGGCGGAGGCTCGGAGCTCGAGCGCGCAGCTGAGCTGAGCTGAGCTGAG 584
QY 2343 TCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2402
Db 583 TCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 524
QY 2403 CCGAGTCCCATGCAACCAAGGAGGCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 2462
Db 523 CCGAGTCCCATGCAACCAAGGAGGCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 465
QY 2463 GGCAGCTACCCCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2522
Db 464 GGCAGCTACCCCTGAGCCCTGAGAGTGGCGGCGCAGCGGAGCGGAGCTGCA 405
QY 2523 GTCGTGGAGACTTGGAGAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2572
Db 404 GTCGTGGAGACTTGGAGAACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
QY 2573 -----ACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2601
Db 344 GCATTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
QY 2602 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2661
Db 284 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225
QY 2662 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2712
Db 224 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 165
QY 2713 CTAGCTCAAGGCTTGTAAACACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTG 2772
Db 164 CTAGCTCAAGGCTTGTAAACACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTG 108
QY 2773 TGCACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2798
Db 107 TGCACCAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 82

PRIOR APPLICATION NUMBER: 06/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 06/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent version 3.0
; SEQ ID NO 40
; LENGTH: 161652
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-10-081-327-40

Query Match 10.4%; Score 647.2; DB 15; Length 161652;
Best Local Similarity 62.3%; Pred. No. 5.8e-160;
Matches 1303; Conservat 0; Mismatches 678; Indels 109; Gaps 14;

DY 780 TAAATATCTTTTAAATTTTCCCTTAACCTTTTAAAGTAAAGTAAAGCCGCGCGG 839
Db 139002 TAAATATCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 139061
QY 840 CCGCATGTCAGCGCTGTAAATTCAGCACTTGGAGGCTGAGGCGGAGATCACTTGA 899
Db 139062 GAGGTGGGCACTGCTGTATTCATGACACTTTGGAGCCCAAGTGGGAATGACTTGA 139121
QY 900 GATCAACAGTTCGAGACCACTGCGCCACATGACAAACCCCATTTTCTACTAAATTA 959
Db 139122 GCTCAAGATTTTGAGACCACTGCGCCCAATGATGACACC-----TCTCAAAATTTT 139176
QY 960 AAAAATGATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019
Db 139177 AAAAATGATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 139236
QY 1020 AGGAAATGCTTGAACCTGGAAGCGGAGGTTGCACTGAGCCCAATCATGCTGCA 1079
Db 139237 GGGAAATGCTTGAACCTGGAAGCGGAGGTTGCACTGAGCCCAATCATGCTGCA 139296
QY 1080 CTCAGCCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1139
Db 139297 CTCAGCCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 139354
QY 1140 TTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
Db 139355 AACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 139414
QY 1200 TACCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
Db 139415 T-----TACAGCAGCTGTAAGAGGCTGATTAATTCATTTATATATGATGA 139464
QY 1260 TACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
Db 139465 ATGTGAG-----GCAGACAGATTTAAAGAAATTTAAAGCCAGGCACTGCT 139510
QY 1320 TGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
Db 139511 GGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 139570
QY 1380 AAGTGGAGGAGGAGCTGCGACGCCCTTCTAACCATGAAAGAAATACCTGCTAGAGCTTC 1439
Db 139571 AGATCGAGACCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 139630
QY 1440 TGGATGCTGGAAGATGATTAACGGGGGCTTCTGAGAGCTGCGCCCTGCTGATGATC- 1498
Db 139631 AAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 139680
QY 1499 TGACTCTGAGCCTCCAGTCAGCTCAGGCGCCATGCTGATGATGATGATGATGATGATG 1558
Db 139691 GCGAGGCTGAGTGAAGGAGCTGCGCTTAACCCGAGAGCGAGAGTGGCACTTGAAGCA 139730
QY 1559 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1618
Db 139731 TCGAGCAGCTG-----CACTCAGAGCTGGGCGGACAG--AGCAAGACTGCGCTC 139795
QY 1619 ATTCAAGATGTACAGCTTTCTTTCAGAGGAAAGATGATGATGATGATGATGATGATG 1678

Db 6911 GTGTGAGAGGAAAGGCGCGAGGCGGAACCGGGGCTGGCGCGGCGCTTGGCGGCGCCCTG 6852
Qy 2107 GAGTTCGGGCTGGGCGCTGGGCTTGGGCGGCGCCGCACTCGAGAGCGGGGCCACCTGTC 2166
Db 6851 GAGTTCGGGCTGGGCGCTGGGCTTGGGCGGCGCCGCACTCGAGAGCGGGGCCCGGCGC -TG 6794
Qy 2167 CAGGCGCGCGGCAATGAGAGCTTAGACCGCGGCGCGGCGCGGCGGCGGAGGCTGACGCGG 2226
Db 6793 CCGGCGCGGCAATGAGAGGCTTAGACCGCGGCGCGGCGCGGCGGCGGAGGCTGACGCGG 6734
Qy 2227 TGGCCCAAGAGTGGCAAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2286
Db 6733 TCCCGCAAGAGTGGCAAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6678
Qy 2287 AGCGTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2341
Db 6677 TGGCTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6618
Qy 2342 CTGCATGGGCTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2401
Db 6617 CTGCATGGGCTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6558
Qy 2402 GCCCAGTCCCATGAGACCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2461
Db 6557 GCCCAGTCCCATGAGACCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6499
Qy 2462 AGCAGCTACCCCTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2521
Db 6498 AGCAGCTACCCCTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6439
Qy 2522 AGCTGTGGAGAGCTTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2581
Db 6438 AGCTGTGGAGAGCTTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6379
Qy 2582 AGCAGCTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2641
Db 6378 AGCAGCTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6319
Qy 2642 CTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2692
Db 6318 CTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6259
Qy 2693 CACCAATCGGCACTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2751
Db 6258 GAGGTGAGAAATTTGTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6200

RESULT 13

US-10-081-327-38/c
; Sequence 38, Application US/10081327
; Publication No. US20030129602A1
; GENERAL INFORMATION:
; APPLICANT: Huang, T1m
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/10/081,327
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 06/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 06/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 128779
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-10-081-327-38

Query Match 9.48; Score 588.6; DB 15; Length 128779;
Best Local Similarity 83.28; Pred. No. 1,9e-144;
Matches 778; Conservative 0; Mismatches 94; Indels 63; Gaps 7;

Qy 1873 TGAAGGTGACAGCATCCCGGCACTCTCACAGCCCTCTGTTCCGCTTCGGGCGCTCTCT 1932
Db 88576 TGAAGGTGACAGCATCTCTGCACTCTCTCACAGCCCTCTGTTCCGCTTCGGGCGCTCTCT 88517
Qy 1933 GCCGCGGCTCCCACTTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1992
Db 88516 GCCGCGGCTCCCACTTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88457
Qy 1993 AGCCCTTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2052
Db 88456 AGCCCTTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88397
Qy 2053 AGGAGAGGCTCAAGCAGAAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2112
Db 88396 AGGAGAGGCTCAAGCAGAAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88337
Qy 2113 CGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2172
Db 88336 CGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88278
Qy 2173 CCGGCGGCAATGAGAGGCTTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2232
Db 88277 CCGGCGGCAATGAGAGGCTTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88218
Qy 2233 AGCAGTCCAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2292
Db 88217 AGCAGTCCAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88162
Qy 2293 CCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2347
Db 88161 CCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 88103
Qy 2348 GGGCTCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2407
Db 88102 GGGCTCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 88043
Qy 2408 TCCCATGACACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2467
Db 88042 TCCCATGACACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 87984
Qy 2468 CTACCCCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2527
Db 87983 CTACCCCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 87924
Qy 2528 GTGGAGACTTGGAGAACCTTATGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2572
Db 87923 GTGGAGACTTGGAGAACCTTATGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 87864
Qy 2573 -----ACACCAATGACAGCCCTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2606
Db 87863 CTGTATCTAGCTCAAGGTTGTAAACACCAATGACAGCCCTGTGTAGTGTAGTGTAGTGTAGT 87804
Qy 2607 TGTGAATGACACCAATGACACCTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 2666
Db 87803 TGTGAATGACACCAATGACACCTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 87744
Qy 2667 TATGTAGTCAAGGATGTGTAAATACCAATGCGCA-----CTGTGTATCTAG 2716
Db 87743 TGTGTGAATCTGTGTATCTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 87684
Qy 2717 CTCAAGGTTGTAAACACCAATGACAGCCCTGT 2751
Db 87683 CTCAAGGATGTAAACGACCAATGACAGCCCTGT 87649

RESULT 14

US-09-764-847-1878/c
; Sequence 1878, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, proteins, and antibodies
; FILE REFERENCE: PC009

; CURRENT APPLICATION NUMBER: US/09/764, 847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1878
; LENGTH: 6766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1878

Query Match 9.4%; Score 583.8; DB 10; Length 6766;
Best Local Similarity 85.3%; Pred. No. 5.8e-144;
Matches 719; Conservative 0; Mismatches 72; Indels 52; Gaps 4;

QY 1873 TGAAGGTGACAGATGCCGCGACTCTCAGACGCTTGTGCTGCGGCGCTCTCT 1932
DB 4597 TGAGAGGTGACAGGCTGTGCTGACCTCAGACGCTTGTGCTGCGGCGCTCTCT 4538
QY 1933 GCGTGGGCTCCACTTGGTGGGCACTTGAGGAGGCGCTTCAAGCCGCTGACCTGTTGG 1992
DB 4537 GCGTGGGCTCCACTTGGTGGGCACTTGAGGAGGCGCTTCAAGCCGCTGACCTGTTGG 4478
QY 1993 AGCCCTTTCTGGGCTGGCCAAAGGCGCAGAGCGGCTCCCTCAGCTTGCAGGAGGTGTGG 2052
DB 4477 AGCCCTTTCTGGGCTGGCCAAAGGCGCAGAGCGGCTCCCTCAGCTTGAAGGAGGTGTGG 4418
QY 2053 AGGAGAGGCTCAGACAGGAGAACCGGGGCTGGGCAAGCGGCTTGGGGGCGCTGAGTTTC 2112
DB 4417 AGGAGAGGCTCAGACAGGAGAACCGGGGCTGGGCAAGCGGCTTGGGGGCGCTGAGTTTC 4358
QY 2113 GCGTGGGCTGGGCTTGGGCGGGGCGCCGCACTCGGAGAGGCGGCGCAGCCCTGCGAGGCC 2172
DB 4357 GCGTGGGCTGGGCTTGGGCGGGGCGCCGCAACTCAGAGAGCGGCTGCTGCTGCC--GCC 4300
QY 2173 CCGGGCAATGAGAGGCTTGAACCCGCGGCGCAGCGGCTGGAGGCTGTACTGGGTGCC 2232
DB 4299 CCGGGCAATGAGAGGCTTGAACCCGCGGCGCAGCGGCTGGAGGCTGTACTGGGTGCC 4240
QY 2233 AGCAGTGGCAGCGGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
DB 4239 AGAAGTGGCAGCGGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4184
QY 2293 CCGCGGGGCGAGGGGCTGGGCACTGACGCGGCGGCGGCTGAGGCT-----CCGCTCAT 2347
DB 4183 CCGCGGGGCGAGGGGCTGGGCACTGACGCGGCGGCGGCTGAGGCTTCCACGCTCCCTGCT 4124
QY 2348 GGGCTCTGTGCGGCGGCGGCGGCTCCGAGAGGCGCAGCAGCCGCTGCTCAGAGGCGCCAG 2407
DB 4123 GGGCTCTGTGCGGCGGCGGCGGCTCCGAGAGGCGCAGCAGCCGCTGCTCAGAGGCGCCAG 4064
QY 2408 TCCCATGAGCAGCAGGAGGCTGAGAGGCTGGGCGGCGCAGGCGCAGCGGAGCTGGAGCGAG 2467
DB 4063 TCCCATGAGCAGCAGGAGGCTGAGAGGCTGGGCGGCGCAGGCGCAGCGGAGCTGGAGCGAG 4004
QY 2468 CTACCCCTGAGAGGCTTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2527
DB 4003 CTACCCCTGAGAGGCTTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3944
QY 2528 GTGAGAGCTTGGAGAGCTTATGTATGTCTAGCTCAGGGATCGTAATATACCAATAGCAGCC 2587
DB 3943 GTGAGAGCTTGGAGAGCTTATGTATGTCTAGCTCAGGGATCGTAATATACCAATAGCAGCC 3922
QY 2588 CTGTGTCTAGCTCAGGGCTGTGTAATGACCAATTCACACTCTGTATCTAGCTACTGTGA 2647
DB 3921 ---GTCTAGCTCAGGGCTGTGTAATGACCAATTCACACTCTGTATCTAGCTACTGTGA 3865
QY 2648 TGGGGCTTGGAGAGCTTATGTATGTCTAGCTCAGGGATGTAAATATACCAATTCGGCAGCTC 2707
DB 3864 TGGGGCAATGGAGAGCTTGTGTGTAGCTCAGGGATGTAAAGCAGCAACCAAGCGGCC 3805
QY 2708 TGT 2710
DB 111

DB 3804 TGT 3802

RESULT 15
US-10-092-154-1878/c
; Sequence 1878; Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC00901
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1878
; LENGTH: 6766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1878

Query Match 9.4%; Score 583.8; DB 14; Length 6766;
Best Local Similarity 85.3%; Pred. No. 5.8e-144;
Matches 719; Conservative 0; Mismatches 72; Indels 52; Gaps 4;

QY 1873 TGAAGGTGACAGATGCCGCGACTCTCAGACGCTTGTGCTGCGGCGCTCTCTCT 1932
DB 4597 TGAGAGGTGACAGGCTGTGCTGACCTCAGACGCTTGTGCTGCGGCGCTCTCTCT 4538
QY 1933 GCGTGGGCTCCACTTGGTGGGCACTTGAGGAGGCGCTTCAAGCCGCTGACCTGTTGG 1992
DB 4537 GCGTGGGCTCCACTTGGTGGGCACTTGAGGAGGCGCTTCAAGCCGCTGACCTGTTGG 4478
QY 1993 AGCCCTTTCTGGGCTGGCCAAAGGCGCAGAGCGGCTCCCTCAGCTTGCAGGAGGTGTGG 2052
DB 4477 AGCCCTTTCTGGGCTGGCCAAAGGCGCAGAGCGGCTCCCTCAGCTTGAAGGAGGTGTGG 4418
QY 2053 AGGAGAGGCTCAGACAGGAGAACCGGGGCTGGGCAAGCGGCTTGGGGGCGCTGAGTTTC 2112
DB 4417 AGGAGAGGCTCAGACAGGAGAACCGGGGCTGGGCAAGCGGCTTGGGGGCGCTGAGTTTC 4358
QY 2113 GCGTGGGCTGGGCTTGGGCGGGGCGCCGCACTCGGAGAGGCGGCGCAGCCCTGCGAGGCC 2172
DB 4357 GCGTGGGCTGGGCTTGGGCGGGGCGCCGCAACTCAGAGAGCGGCTGCTGCTGCC--GCC 4300
QY 2173 CCGGGCAATGAGAGGCTTGAACCCGCGGCGCAGCGGCTGGAGGCTGTACTGGGTGCC 2232
DB 4299 CCGGGCAATGAGAGGCTTGAACCCGCGGCGCAGCGGCTGGAGGCTGTACTGGGTGCC 4240
QY 2233 AGCAGTGGCAGCGGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
DB 4239 AGAAGTGGCAGCGGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4184
QY 2293 CCGCGGGGCGAGGGGCTGGGCACTGACGCGGCGGCGGCTGAGGCT-----CCGCTCAT 2347
DB 4183 CCGCGGGGCGAGGGGCTGGGCACTGACGCGGCGGCGGCTGAGGCTTCCACGCTCCCTGCT 4124
QY 2348 GGGCTCTGTGCGGCGGCGGCGGCTCCGAGAGGCGCAGCAGCCGCTGCTCAGAGGCGCCAG 2407
DB 4123 GGGCTCTGTGCGGCGGCGGCGGCTCCGAGAGGCGCAGCAGCCGCTGCTCAGAGGCGCCAG 4064
QY 2408 TCCCATGAGCAGCAGGAGGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2467
DB 4063 TCCCATGAGCAGCAGGAGGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4004
QY 2468 CTACCCCTGAGAGGCTTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2527
DB 4003 CTACCCCTGAGAGGCTTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3944
QY 2528 GTGAGAGCTTGGAGAGCTTATGTATGTCTAGCTCAGGGATCGTAATATACCAATAGCAGCC 2587
DB 3943 GTGAGAGCTTGGAGAGCTTATGTATGTCTAGCTCAGGGATCGTAATATACCAATAGCAGCC 3922

Qy	2588	CTGTCTAGCTCAGGCTCTGTGATGACCAATCCACACTCTGTATCTAGTACTCTGA	2647
Db	3921	---TGCTAGCTCAGGCTTGTGATGACACAGTGACACTGTATCTAGCTAATCTGG	3865
Qy	2648	TGGGGCCTTGAGAACCTTATGCTAGCTCAGGATTTGAATACACCAATCGGCACTC	2707
Db	3864	TGGGGACATGAGAACCTTGTGCTAGCTCAGGATTTGAACGACCAACCAAGCGGCC	3805
Qy	2708	TGT 2710	
Db	3804	TGT 3802	

Search completed: August 17, 2003, 01:54:18
Job time : 1003 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2003, 13:22:16 ; Search time 11294 Seconds
(without alignments)
18089.916 Million cell updates/sec

Title: US-09-845-020A-5
Perfect score: 6235
Sequence: 1 gataccttgagagacagtagt.....tgcacattccccacagcgct 6235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

- 1: /cgn2_6/ptodata/2/pna/PCrus.COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/PCrus.COMB.seq.old.*
- 3: /cgn2_6/ptodata/2/pna/US06.COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US07.COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US08.COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US082.COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US083.COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US084.COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US085.COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US086.COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US087.COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US088.COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US089.COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US090.COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US091.COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092.COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093.COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US094.COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US095.COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US096.COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US097.COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US098.COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US099.COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US100.COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US101.COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US102.COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US103.COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US104.COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US105.COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US106.COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US107.COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US108.COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US109.COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US110.COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US111.COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US112.COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US113.COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US114.COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US115.COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US116.COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US117.COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US118.COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US119.COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US120.COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US121.COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US122.COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US123.COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US124.COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US125.COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US126.COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US127.COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US128.COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US129.COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US130.COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US131.COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US132.COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US133.COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US134.COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US135.COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US136.COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US137.COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US138.COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US139.COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US140.COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US141.COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US142.COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US143.COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US144.COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US145.COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US146.COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US147.COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US148.COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US149.COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US150.COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US151.COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US152.COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US153.COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US154.COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US155.COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US156.COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US157.COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US158.COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US159.COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US160.COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US161.COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US162.COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US163.COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US164.COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US165.COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US166.COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US167.COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US168.COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US169.COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US170.COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US171.COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US172.COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US173.COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US174.COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US175.COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US176.COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US177.COMB.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6235	100.0	6235	35	US-09-845-020A-5
2	6235	100.0	6679	35	US-09-845-020A-1
3	6056	97.1	4925599	40	US-09-948-128-259
4	6052.8	97.1	14371	101	US-60-466-412-83872

5	5532.6	88.7	14381	101	US-60-465-241-55636	Sequence 52636, A
6	5532.6	88.7	14381	101	US-60-466-412-86425	Sequence 86425, A
7	2834	45.5	2834	35	US-09-845-020A-6	Sequence 6, App1
8	2225.8	35.7	2409	33	US-09-760-457-490	Sequence 490, App
9	2225.8	35.7	2409	33	US-09-760-463-27	Sequence 27, App1
10	2225.8	35.7	2409	33	US-09-760-485-1047	Sequence 1047, Ap
11	2225.8	35.7	2409	46	US-10-143-877-227	Sequence 27, App1
12	2225.8	35.7	2409	48	US-10-216-436-1047	Sequence 1047, Ap
13	2225.8	35.7	2409	48	US-10-217-527-492	Sequence 490, App
14	2225.8	35.7	2409	48	US-10-217-651-592	Sequence 592, App
15	1959.4	31.4	46997	101	US-60-466-412-86428	Sequence 86428, A
16	1252	20.1	1252	35	US-09-845-020A-7	Sequence 7, App1
17	1050.2	16.8	4363157	40	US-09-947-911-65	Sequence 65, App1
18	1037	16.8	3037327	40	US-09-947-911-286	Sequence 286, App
19	1025.8	16.5	6853926	40	US-09-947-916-137	Sequence 137, App
20	1025.6	16.4	59140	40	US-09-949-016-16623	Sequence 16023, A
21	1025.6	16.4	67140	101	US-60-466-412-86762	Sequence 86762, A
22	1002.2	16.1	36221	1	PCT-US02-29149-29	Sequence 29, App1
23	1002.2	16.1	36221	41	US-09-954-556-29	Sequence 29, App1
24	986.2	15.8	59160	79	US-60-245-241-16	Sequence 16, App1
25	972	15.6	185759	76	US-60-226-176-146	Sequence 16, App1
26	972	15.6	185759	77	US-60-233-468-1446	Sequence 1446, Ap
27	972	15.6	185759	86	US-60-213-371-1446	Sequence 1446, Ap
28	956.8	15.3	32768	75	US-60-313-178-377	Sequence 1446, Ap
29	956.6	15.3	540438	40	US-09-947-911-332	Sequence 377, App
30	944.2	15.1	3037327	40	US-09-947-911-286	Sequence 322, App
31	940.2	15.1	8616041	40	US-09-947-916-174	Sequence 286, App
32	924.2	14.8	1270403	40	US-09-947-911-221	Sequence 174, App
33	923.8	14.8	32768	77	US-60-233-940-9	Sequence 221, App
34	917.4	14.7	89635	45	US-10-087-119-2020	Sequence 9, App1
35	917.4	14.7	177558	99	US-60-449-629-846	Sequence 2020, Ap
36	915.4	14.7	246320	40	US-09-949-016-17019	Sequence 846, App
37	915.4	14.7	246320	40	US-09-949-016-17019	Sequence 17019, A
38	915.4	14.7	246320	40	US-09-949-016-17020	Sequence 17020, A
39	915.4	14.7	246320	40	US-09-949-016-17021	Sequence 17021, A
40	896.4	14.4	49899	101	US-60-466-412-84731	Sequence 17022, A
41	890.2	14.3	1044426	40	US-09-947-911-240	Sequence 84731, A
42	890	14.3	185759	76	US-60-226-176-146	Sequence 240, App
43	890	14.3	185759	77	US-60-233-466-1446	Sequence 1446, Ap
44	890	14.3	185759	86	US-60-313-371-1446	Sequence 1446, Ap
45	880.8	14.1	499202	101	US-60-465-241-51856	Sequence 51856, A

1081 TCCAGCTGGGTGACAGAGTGAACCTTGTCTCAAGCAAAAAAAGTGTAAAGGCAT 1140
1081 TCCAGCTGGGTGACAGAGTGAACCTTGTCTCAAGCAAAAAAAGTGTAAAGGCAT 1140
1141 TCCATTAATTCAGTGTACATCAGTGTACATCTACAGTGTGGGTACTCTGCTGTAGGCAT 1200
1141 TCCATTAATTCAGTGTACATCAGTGTACATCTACAGTGTGGGTACTCTGCTGTAGGCAT 1200
1201 ACCGTGAAATAGTAGTGTCTTGGTCAACAGACATACATCTTCCACATTTAATGTAGACT 1260
1201 ACCGTGAAATAGTAGTGTCTTGGTCAACAGACATACATCTTCCACATTTAATGTAGACT 1260
1261 ACCAAGTTCATCCAGAGAGGTTTCTTTTACAAATCTACATCTCCCGCAGCAAAAT 1320
1261 ACCAAGTTCATCCAGAGAGGTTTCTTTTACAAATCTACATCTCCCGCAGCAAAAT 1320
1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAAACAGA 1380
1321 GAGAGTTACTCCAGATCTTTTACAAAGATGCTTAAAGCCAGTACAGATGAAAACAGA 1380
1381 AGTGGGAGGGAAGCTGCCAGCCCTCTTAACCATGAAAGAAATACCTGTGTAGAGCTTCT 1440
1381 AGTGGGAGGGAAGCTGCCAGCCCTCTTAACCATGAAAGAAATACCTGTGTAGAGCTTCT 1440
1441 GATGTCTGAAGATGAATAAACGGGGGTCTGTGGAGCTGCCCTGTCAAGATCACTGTG 1500
1441 GATGTCTGAAGATGAATAAACGGGGGTCTGTGGAGCTGCCCTGTCAAGATCACTGTG 1500
1501 ACTTGTAGCCCTCCAGTCCAGTCTCAAGCCCATGTGTATGGCCAGTGAATATAGGCCT 1560
1501 ACTTGTAGCCCTCCAGTCCAGTCTCAAGCCCATGTGTATGGCCAGTGAATATAGGCCT 1560
1561 CACTCTGTGTGTCTTATATCTCCCATGTGGGGTGAAGTGTGAAGTGAAGCTTAT 1620
1561 CACTCTGTGTGTCTTATATCTCCCATGTGGGGTGAAGTGTGAAGTGAAGCTTAT 1620
1621 TCAAGATGTACAGCTTCTTGAAGAAATAGTGTACAGAAACACAGAGGGCTTGGCA 1680
1621 TCAAGATGTACAGCTTCTTGAAGAAATAGTGTACAGAAACACAGAGGGCTTGGCA 1680
1681 AGATGATCTAATCTCAATCTTACCTGTGTGACAGCCAGCTGTGTGTGTGTGTGTGT 1740
1681 AGATGATCTAATCTCAATCTTACCTGTGTGACAGCCAGCTGTGTGTGTGTGTGTGT 1740
1681 AGATGATCTAATCTCAATCTTACCTGTGTGACAGCCAGCTGTGTGTGTGTGTGTGT 1740
1681 AGATGATCTAATCTCAATCTTACCTGTGTGACAGCCAGCTGTGTGTGTGTGTGTGT 1740
1741 AAGTTTTTCACTTCTGTGAGGCATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
1741 AAGTTTTTCACTTCTGTGAGGCATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
1741 AAGTTTTTCACTTCTGTGAGGCATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
1801 TGAATGACAGAGTCCCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
1801 TGAATGACAGAGTCCCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
1801 TGAATGACAGAGTCCCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
1801 TGAATGACAGAGTCCCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
1861 ATGGCTTGAACCTGTGAGAGTGAACAGATGCGGAGAGTCTACAGCCCTGTGTGTGTGT 1920
1861 ATGGCTTGAACCTGTGAGAGTGAACAGATGCGGAGAGTCTACAGCCCTGTGTGTGTGT 1920
1921 GGGGCTCCCTGT 1980
1921 GGGGCTCCCTGT 1980
1921 GGGGCTCCCTGT 1980
1981 CTGCACTGT 2040
1981 CTGCACTGT 2040
1981 CTGCACTGT 2040
2041 AGGAGAGT 2100
2041 AGGAGAGT 2100
2101 CAGCTGT 2160
2101 CAGCTGT 2160
2161 CCCTGT 2220
2161 CCCTGT 2220
2221 ACTGGGTGCCAGACAGT 2280
2221 ACTGGGTGCCAGACAGT 2280
2221 ACTGGGTGCCAGACAGT 2280
2281 CTTAGCAGCTTCCCGGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
2281 CTTAGCAGCTTCCCGGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
2281 CTTAGCAGCTTCCCGGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
2341 CTTAGCAGCTTCCCGGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
2341 CTTAGCAGCTTCCCGGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
2401 CGCCAGTCCCATGACAGCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
2401 CGCCAGTCCCATGACAGCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
2461 CAGGCACTACCCCTGT 2520
2461 CAGGCACTACCCCTGT 2520
2521 GAGT 2580
2521 GAGT 2580
2521 GAGT 2580
2581 CAGCACCCTGT 2640
2581 CAGCACCCTGT 2640
2641 ACTGT 2700
2641 ACTGT 2700
2641 ACTGT 2700
2701 GGCACCTGT 2760
2701 GGCACCTGT 2760
2701 GGCACCTGT 2760
2761 AGGATGT 2820
2761 AGGATGT 2820
2761 AGGATGT 2820
2821 GAAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2880
2821 GAAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2880
2881 GGGCTGT 2940
2881 GGGCTGT 2940
2881 GGGCTGT 2940
2881 GGGCTGT 2940
2941 TTGGGTAGTAAAGAAAAATTTACAGTCAAGAGGGGGTGTGTGTGTGTGTGTGTGTGT 3000
2941 TTGGGTAGTAAAGAAAAATTTACAGTCAAGAGGGGGTGTGTGTGTGTGTGTGTGTGT 3000
2941 TTGGGTAGTAAAGAAAAATTTACAGTCAAGAGGGGGTGTGTGTGTGTGTGTGTGTGT 3000
2941 TTGGGTAGTAAAGAAAAATTTACAGTCAAGAGGGGGTGTGTGTGTGTGTGTGTGTGT 3000
3001 GGGGTGCCAAGGT 3060
3001 GGGGTGCCAAGGT 3060
3061 TTTCAGAGTATGTATCAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
3061 TTTCAGAGTATGTATCAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
3061 TTTCAGAGTATGTATCAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
3121 GAAATGTATCAATTAAGT 3180
3121 GAAATGTATCAATTAAGT 3180
3181 TTTCAGAGTATGTATCAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
3181 TTTCAGAGTATGTATCAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
3181 TTTCAGAGTATGTATCAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
3241 CAGAGCTGTACAGCTCTGT 3300
3241 CAGAGCTGTACAGCTCTGT 3300

Db 3241 CAGAGGCTTGACACTACTGTGGTGGGCTTGGAGAAATGTTTGTGTGCACACTCTGTAT 3300
OY 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTGTGTAGCTCAGAGGATTTGTAACGCA 3360
Db 3301 CTAGTTAATCTAGTGGGAGCTGGAGAACCTTGTGTAGCTCAGAGGATTTGTAACGCA 3360
OY 3361 CCAATCAGGCGCTGTCAAAACAGACACTCGGCTCTACCAATCAGCAGAGTGTGGTGG 3420
Db 3361 CCAATCAGGCGCTGTCAAAACAGACACTCGGCTCTACCAATCAGCAGAGTGTGGTGG 3420
OY 3421 GGGCAGATAGAGAAATAAAAGCAGGCTGGCCGAGCAGAGTGGCAACGGCAGAGTGC 3480
Db 3421 GGGCAGATAGAGAAATAAAAGCAGGCTGGCCGAGCAGAGTGGCAACGGCAGAGTGC 3480
OY 3481 CTATCCACAATATGAGAGCTTGTGTCTTGTGTGGATTAATCTTGCTACTGCTCG 3540
Db 3481 CTATCCACAATATGAGAGCTTGTGTCTTGTGTGGATTAATCTTGCTACTGCTCG 3540
OY 3541 CTTTTTGGGTCCACACTGTGTTTATGAGCTGTAACTCAGCAGAAAGTCTGACACTTC 3600
Db 3541 CTTTTTGGGTCCACACTGTGTTTATGAGCTGTAACTCAGCAGAAAGTCTGACACTTC 3600
OY 3601 ACTCCTGAAGCCACTTAAGACCAACAGCCACCGGAGGAATGAACAATCCGGCCGCT 3660
Db 3601 ACTCCTGAAGCCACTTAAGACCAACAGCCACCGGAGGAATGAACAATCCGGCCGCT 3660
OY 3661 GCCTTAAGAGCTTAACACTCACCGGAAGTGTGCAGCTTCACTCTCAGCCAGCGAGA 3720
Db 3661 GCCTTAAGAGCTTAACACTCACCGGAAGTGTGCAGCTTCACTCTCAGCCAGCGAGA 3720
OY 3721 CCACGAACCCACAGAGGAAGAACTGGGAACACTGTGAACATAGAAAGAAACAAC 3780
Db 3721 CCACGAACCCACAGAGGAAGAACTGGGAACACTGTGAACATAGAAAGAAACAAC 3780
OY 3781 CCAGATGCACCACTTAAGAGCTGTAACTCACTCAGAGGTCGCGGCTTCTCTTG 3840
Db 3781 CCAGATGCACCACTTAAGAGCTGTAACTCACTCAGAGGTCGCGGCTTCTCTTG 3840
OY 3841 AAGTCAGTGAAGCAACAGCACTCACAGTTTCGGACAAGCCAGAGATTGAGATCAG 3900
Db 3841 AAGTCAGTGAAGCAACAGCACTCACAGTTTCGGACAAGCCAGAGATTGAGATCAG 3900
OY 3901 CTGGGCAACATGATGAATAGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCG 3960
Db 3901 CTGGGCAACATGATGAATAGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCG 3960
OY 3961 AGCATGTTGGTCCGTGCTGTGGTCCAGCTACGCGGAGGCTTAAAGTGGAGATCGCT 4020
Db 3961 AGCATGTTGGTCCGTGCTGTGGTCCAGCTACGCGGAGGCTTAAAGTGGAGATCGCT 4020
OY 4021 TGAAGCTGGAGGTGAAGACTGCACTGAGCTGTGATTTGACACAGCCCTTAAGCTGG 4080
Db 4021 TGAAGCTGGAGGTGAAGACTGCACTGAGCTGTGATTTGACACAGCCCTTAAGCTGG 4080
OY 4081 GGAACAGCTGAGAGCCCTGTTCCCTCCCAAAAAAATTTGACAAAGTGAATAAAGGT 4140
Db 4081 GGAACAGCTGAGAGCCCTGTTCCCTCCCAAAAAAATTTGACAAAGTGAATAAAGGT 4140
OY 4141 GCGTGTATGCTAGAGCGAGTGGCTCATGCTGTAACTCCAGCACTTTGGAGCGGAG 4200
Db 4141 GCGTGTATGCTAGAGCGAGTGGCTCATGCTGTAACTCCAGCACTTTGGAGCGGAG 4200
OY 4201 GCGGCGGGGTCACTTAAGTCAAGAGTGTGAACAGCAGCTGGCCAACTGGAGAAAGCCC 4260
Db 4201 GCGGCGGGGTCACTTAAGTCAAGAGTGTGAACAGCAGCTGGCCAACTGGAGAAAGCCC 4260
OY 4261 ATCTCTTTAAAAAATTAAGTGAAGTGTGGGAGGAGTGTGGAGATGCTGTAA 4320
Db 4261 ATCTCTTTAAAAAATTAAGTGAAGTGTGGGAGGAGTGTGGAGATGCTGTAA 4320
OY 4321 TCCAGACTACTAGAGAGGCTGAAGCGAGGAATCACTTGAACCCAGAGAGGCGGCTGC 4380
Db 4321 TCCAGACTACTAGAGAGGCTGAAGCGAGGAATCACTTGAACCCAGAGAGGCGGCTGC 4380

OY 4381 AGTAGCCGAGATCGTGCCATTGCTCACTCCACCACCTCAGGCTGGGCAACAGAGCCAAA 4440
Db 4381 AGTAGCCGAGATCGTGCCATTGCTCACTCCACCACCTCAGGCTGGGCAACAGAGCCAAA 4440
OY 4441 CTCTGCTTTAAAAAATTAAGTGTGAGTGTGAATGAAGTGTGCAATGCAATAG 4500
Db 4441 CTCTGCTTTAAAAAATTAAGTGTGAGTGTGAATGAAGTGTGCAATGCAATAG 4500
OY 4501 TTGCGAGGCAACATGTTTAAAGATGTGAGTCTCTCTCCATGGTCTGTTAAAAAC 4560
Db 4501 TTGCGAGGCAACATGTTTAAAGATGTGAGTCTCTCTCCATGGTCTGTTAAAAAC 4560
OY 4561 CACCTCAAGGCGAGGTGAGTGTCAATGCTATTAATCCAGCACTTTGGAGGCGGAG 4620
Db 4561 CACCTCAAGGCGAGGTGAGTGTCAATGCTATTAATCCAGCACTTTGGAGGCGGAG 4620
OY 4621 GCGGAGTGAATCACTGAGTCAAGTGTGAAGCCAGCCCTGACCAACATGTTAAAT 4680
Db 4621 GCGGAGTGAATCACTGAGTCAAGTGTGAAGCCAGCCCTGACCAACATGTTAAAT 4680
OY 4681 CCCACCTCTACTAAAAATTAAGATGAGCATGTGTGTCATGCTGTAAATCCAC 4740
Db 4681 CCCACCTCTACTAAAAATTAAGATGAGCATGTGTGTCATGCTGTAAATCCAC 4740
OY 4741 CTACTTGGAGGCTGAGGCGAGAAATCACTAGAACAGGAGGCGGAGTGTACTGAG 4800
Db 4741 CTACTTGGAGGCTGAGGCGAGAAATCACTAGAACAGGAGGCGGAGTGTACTGAG 4800
OY 4801 CCGAGATCGTGCCATTTGCACTCCAGCCTGAGCAATGAGGAATCTCATCTCAAAAAA 4860
Db 4801 CCGAGATCGTGCCATTTGCACTCCAGCCTGAGCAATGAGGAATCTCATCTCAAAAAA 4860
OY 4861 AACAAAAAACCACACTCTCTACTCCAGGAGACTGGGTACAGAGCTGGGCCACATCAG 4920
Db 4861 AACAAAAAACCACACTCTCTACTCCAGGAGACTGGGTACAGAGCTGGGCCACATCAG 4920
OY 4921 GCAAGTGTGAGCCACAGAGCTTAAGCGGAGTGTGAGAGCCGCGACCAAGATTAACGT 4980
Db 4921 GCAAGTGTGAGCCACAGAGCTTAAGCGGAGTGTGAGAGCCGCGACCAAGATTAACGT 4980
OY 4981 TGTGAATCAGTGTGAGATCAAGTCCCTCCCTGTTGTGTGACCAACAGGGGCCCCCA 5040
Db 4981 TGTGAATCAGTGTGAGATCAAGTCCCTCCCTGTTGTGTGACCAACAGGGGCCCCCA 5040
OY 5041 AGCAGCAGATGAGCCCACTCACTCAGCAGCATCACTCTCTCACTCAGAGATGTCTTT 5100
Db 5041 AGCAGCAGATGAGCCCACTCACTCAGCAGCATCACTCTCTCACTCAGAGATGTCTTT 5100
OY 5101 CTTGCGACGCTGGGGTAAATTAAGACAGAGTGTGAGTGTGGTGTGAGTGTGAG 5160
Db 5101 CTTGCGACGCTGGGGTAAATTAAGACAGAGTGTGAGTGTGGTGTGAGTGTGAG 5160
OY 5161 TGGCCGAGGAGGCTTGTGGCTGTGAAGAAAGTTCAGGCCCTAAGGCCGCGGAGTGG 5220
Db 5161 TGGCCGAGGAGGCTTGTGGCTGTGAAGAAAGTTCAGGCCCTAAGGCCGCGGAGTGG 5220
OY 5221 TCAGGCTGTAAATCCAGACATTTGGAGGCGCAGGCGGAGTGTGATCAGAGTCAAGAGA 5280
Db 5221 TCAGGCTGTAAATCCAGACATTTGGAGGCGCAGGCGGAGTGTGATCAGAGTCAAGAGA 5280
OY 5281 TCGTGAACATCTGTGCTAAACAGGTGAAGAACCCGCTCTACTTAAAAATTAAG 5340
Db 5281 TCGTGAACATCTGTGCTAAACAGGTGAAGAACCCGCTCTACTTAAAAATTAAG 5340
OY 5341 CCGGAGCATGTGGCGGAGCACTGTAACTCAGTACTCGGAGGCTGAGGCGAGGAATG 5400
Db 5341 CCGGAGCATGTGGCGGAGCACTGTAACTCAGTACTCGGAGGCTGAGGCGAGGAATG 5400
OY 5401 GCGTGAACCCGAGAGCAGAGTTTGCAGTGAAGCGAGATCGGCGCACTCAGCTCAAGCT 5460
Db 5401 GCGTGAACCCGAGAGCAGAGTTTGCAGTGAAGCGAGATCGGCGCACTCAGCTCAAGCT 5460

QY 5461 GGGCGACAGACAGACTCATCTGAAAAAGAAAAAGCGTTGAGTCTGACAGAGA 5520
DB 5461 GGGCGACAGACAGACTCATCTGAAAAAGAAAAAGCGTTGAGTCTGACAGAGA 5520
QY 5521 GGGCGACAGCTGTAATTTCTGCTACATACCATGACCTTGGGCAAGGACCTTCTGCTGAC 5580
DB 5521 GGGCGACAGCTGTAATTTCTGCTACATACCATGACCTTGGGCAAGGACCTTCTGCTGAC 5580
QY 5581 CCAGTTCACAGGGGTGGGAATGACATCCCAAGGTCCTTCCAGATTAAGCTGCAAGGTTTC 5640
DB 5581 CCAGTTCACAGGGGTGGGAATGACATCCCAAGGTCCTTCCAGATTAAGCTGCAAGGTTTC 5640
QY 5641 TAAGATGAGAAGATGGGAGAGTTTCCCTCTCTACCCAGCCGCTGTCACATTAAGAT 5700
DB 5641 TAAGATGAGAAGATGGGAGAGTTTCCCTCTCTACCCAGCCGCTGTCACATTAAGAT 5700
QY 5701 GAATGACACAGGAGATGACATGTCCTCCCAATCCCGAGTTCCAAAGCCCTTGGGAGCCCTAC 5760
DB 5701 GAATGACACAGGAGATGACATGTCCTCCCAATCCCGAGTTCCAAAGCCCTTGGGAGCCCTAC 5760
QY 5761 TGTGAGGTCGTGACAGAGAGGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5820
DB 5761 TGTGAGGTCGTGACAGAGAGGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5820
QY 5821 CCTCATTTGGGAGACAGACATCCGGTTTCTGCTGCTCTACCGGATTTAGGGGCTTTAGC 5880
DB 5821 CCTCATTTGGGAGACAGACATCCGGTTTCTGCTGCTCTACCGGATTTAGGGGCTTTAGC 5880
QY 5881 CGAATGACTCATGAGGG 5940
DB 5881 CGAATGACTCATGAGGG 5940
QY 5941 CAGGACACCCCTGGAATCTTCATGCTGCTTCCCAAGTGGGGGGGGGGGGGGGGGGGGGGGG 6000
DB 5941 CAGGACACCCCTGGAATCTTCATGCTGCTTCCCAAGTGGGGGGGGGGGGGGGGGGGGGGGG 6000
QY 6001 GACCCCAATGCTTATCTCAGGTAGGGGCTCAGAGAGTCTCCACAGACAGGACCTCCGG 6060
DB 6001 GACCCCAATGCTTATCTCAGGTAGGGGCTCAGAGAGTCTCCACAGACAGGACCTCCGG 6060
QY 6061 AGAGTTGGGGGTAGGAATGGGAGACACAGGCTTCTTTTCTCTCTAGAAATTTGGG 6120
DB 6061 AGAGTTGGGGGTAGGAATGGGAGACACAGGCTTCTTTTCTCTCTAGAAATTTGGG 6120
QY 6121 GGGCTTGGGGGACAGGCTTGAAGATCCCAAGGAGGAGGGGCAAGGACCTCCCGACAG 6180
DB 6121 GGGCTTGGGGGACAGGCTTGAAGATCCCAAGGAGGAGGGGCAAGGACCTCCCGACAG 6180
QY 6181 TCTGCCAGAGGAGAGAGGAGAGACCCGACACTGACCTTCCCAAGGAGGCT 6235
DB 6181 TCTGCCAGAGGAGAGAGGAGAGACCCGACACTGACCTTCCCAAGGAGGCT 6235

RESULT 2
US-09-845-020A-1
Sequence 1, Application us/09845020A

GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heartlein, Michel W.
APPLICANT: Selden, Richard F.
TITLE OF INVENTION: Genomic Sequences for Protein Production
FILE REFERENCE: 50010/017003
CURRENT APPLICATION NUMBER: US/09/845,020A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 09/305,384
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: US 60/084,649
PRIOR FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6679

TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-020A-1
Query Match 100.0%; Score 6235; DB 35; Length 6679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCATTGAGACAGTATGTTAACAACCAAGCTGGGACAGATAGGAGACTCTCTACG 60
DB 20 GATCATTGAGACAGTATGTTAACAACCAAGCTGGGACAGATAGGAGACTCTCTACG 79
QY 61 AAAAATCAAAAAATATATGCGCGGATGAGTGTGCTACGCTGTAATCCCTGAACCTTTGG 120
DB 80 AAAAATCAAAAAATATATGCGCGGATGAGTGTGCTACGCTGTAATCCCTGAACCTTTGG 139
QY 121 ACATCAAGGCAAGTGGATCTTGAAGTTCAGAGTTCGAGTACCTGGCCACATGGT 180
DB 140 ACATCAAGGCAAGTGGATCTTGAAGTTCAGAGTTCGAGTACCTGGCCACATGGT 199
QY 181 GAAACCTATCTCCACTAAAAATATACCAAGCATTTGGTGGACAGCAGCTGTA 240
DB 200 GAAACCTATCTCCACTAAAAATATACCAAGCATTTGGTGGACAGCAGCTGTA 259
QY 241 ATCCGGGCTACTCAGAGGCTGAGGACAGAGAAATCACTTGAACCCAGAGAGGAGGTTG 300
DB 260 ATCCGGGCTACTCAGAGGCTGAGGACAGAGAAATCACTTGAACCCAGAGAGGAGGTTG 319
QY 301 CAGTAGGCTGAGATCAACACACTGCACTCCAGCTGGGTGACAGAGCAAGACTATCTC 360
DB 320 CAGTAGGCTGAGATCAACACACTGCACTCCAGCTGGGTGACAGAGCAAGACTATCTC 379
QY 361 AAAAAAATAAAAAATAAAAAATTTAGCAGGACATGTAAGTGCACACTCTAGTCTAG 420
DB 380 AAAAAAATAAAAAATAAAAAATTTAGCAGGACATGTAAGTGCACACTCTAGTCTAG 439
QY 421 CTACTCAGAGGCTGAGTGGAGGATCACTTGAACCTGAGGAGGACAGGCTAGAGTGA 480
DB 440 CTACTCAGAGGCTGAGTGGAGGATCACTTGAACCTGAGGAGGACAGGCTAGAGTGA 499
QY 481 GCCAAGATCATGCCACTACACTCCAGCTGGGCAACAGAGAGAGACCTGTCTPAAAA 540
DB 500 GCCAAGATCATGCCACTACACTCCAGCTGGGCAACAGAGAGAGACCTGTCTPAAAA 559
QY 541 AAT 600
DB 560 AAT 619
QY 601 ATGTAT 660
DB 620 ATGTAT 679
QY 661 ATGTCTCTGTACATCACTGCTGCTGCTTCTAGACCAAAAGCAGCTATTAACAT 720
DB 680 ATGTCTCTGTACATCACTGCTGCTGCTTCTAGACCAAAAGCAGCTATTAACAT 739
QY 721 ACATACATGAATTTTATATAGATCAGATTTGAATTTCAATATATATATATATATAT 780
DB 740 ACATACATGAATTTTATATAGATCAGATTTGAATTTCAATATATATATATATATAT 799
QY 781 AAAAT 840
DB 800 AAAAT 859
QY 841 GCCATCTGACAGCTGTAATTTCCAGACTTTGGAGGCTGAGTGGGACAGTCACTTGA 900
DB 860 GCCATCTGACAGCTGTAATTTCCAGACTTTGGAGGCTGAGTGGGACAGTCACTTGA 919
QY 901 ATCAACAGTTGGAGACAGGCTGSCCAACATAGCAAAACCCATTTCTACTAAAAATAA 960
DB 920 ATCAACAGTTGGAGACAGGCTGSCCAACATAGCAAAACCCATTTCTACTAAAAATAA 979
QY 961 AAAATAGCTGGGACATAGTGTGACACACTGTGATCCAGCTACTTGGAGGCTAGAGCA 1020

|||||
Db 980 AAAATTAGCTGGCATAGTGTGCACACCTGTGATCCCACTACTTGGAGGCTGAGGCA 1039
QY 1021 GGAGAAATCGCTTGAACCTGGGAAAGCGAGGTTGCAGTGAAGCAACATCATGCCACTGCAC 1080
Db 1040 GGAATAATCGCTTGAACCTGGGAAAGCGAGGTTGCAGTGAAGCAACATCATGCCACTGCAC 1099
QY 1081 TCCAGCCTGGGTGACAGAGTGAAGTCTGTCTCAACGAAAAAAGTGTAAAAAGCAT 1140
Db 1100 TCCAGCCTGGGTGACAGAGTGAAGTCTGTCTCAACGAAAAAAGTGTAAAAAGCAT 1159
QY 1141 TCCAAATTCAGTGTACATCAGTGTACATCTCAGGTGCTGCTACTCCGTCTGAGGCAT 1200
Db 1160 TCCAAATTCAGTGTACATCAGTGTACATCTCAGGTGCTGCTACTCCGTCTGAGGCAT 1219
QY 1201 ACCTGAGAGTGAAGTGTGCTGTGCTACAGAGACATACAACTTCCAACTTAACAGACAT 1260
Db 1220 ACCGAGAGTGAAGTGTGCTGTGCTACAGAGACATACAACTTCCAACTTAACAGACAT 1279
QY 1261 ACCAAGTTCGACCAAGAGAGGTTTTTTTTTAAATCTACACTCCCCCGGCAACAAT 1320
Db 1280 ACCAAGTTCGACCAAGAGAGGTTTTTTTTTAAATCTACACTCCCCCGGCAACAAT 1339
QY 1321 GAGAGTACTCCAGATCTTACAAAGATGCTTAAGCCAGTACAGATGAAGAAACAGA 1380
Db 1340 GAGAGTACTCCAGATCTTACAAAGATGCTTAAGCCAGTACAGATGAAGAAACAGA 1399
QY 1381 AGTGGAGGGAGAGCTGCCAGCCCCCTTCTAAACATGAAGAAATACCTGTGAGACCTTCT 1440
Db 1400 AGTGGAGGGAGAGCTGCCAGCCCCCTTCTAAACATGAAGAAATACCTGTGAGACCTTCT 1459
QY 1441 GGATGCTGGAAGATGAATAACGGGGGTCTGTGAGAGCTGCCCTGTGCAGATACAGTGTG 1500
Db 1460 GGATGCTGGAAGATGAATAACGGGGGTCTGTGAGAGCTGCCCTGTGCAGATACAGTGTG 1519
QY 1501 ACTTCTGAGCCTCCAGTCCAGTCTCAAGCCCATGTGTATGSCCAGTGAATAGGCCCT 1560
Db 1520 ACTTCTGAGCCTCCAGTCCAGTCTCAAGCCCATGTGTATGSCCAGTGAATAGGCCCT 1579
QY 1561 CACTCTGTGTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTGTGATGAGCCGTAT 1620
Db 1580 CACTCTGTGTTGGTCTTTATTTCTCCCATGTGGGGCTGAAGTGTGATGAGCCGTAT 1639
QY 1621 TCAAGATGTACAGCTTCTTGACAGAAAGTGTGACAGAAACACAGAGGCTTGGCA 1680
Db 1640 TCAAGATGTACAGCTTCTTGACAGAAAGTGTGACAGAAACACAGAGGCTTGGCA 1699
QY 1681 AGATGATCTAATGCAATCTTACCTGCTCAGCCACAGCTAGTCTGTGATCTTGAAC 1740
Db 1700 AGATGATCTAATGCAATCTTACCTGCTCAGCCACAGCTAGTCTGTGATCTTGAAC 1759
QY 1741 AAGTTTTTACCTTCTGTGAGGCAATCCCTGGCTACAAACACACAGATTGGTTGACAGA 1800
Db 1760 AAGTTTTTACCTTCTGTGAGGCAATCCCTGGCTACAAACACACAGATTGGTTGACAGA 1819
QY 1801 TGAATATACAAAGTCCCTTACACTGTAAATCCAGCACTTTGGAGGCCAAAGCGGGTGG 1860
Db 1820 TGAATATACAAAGTCCCTTACACTGTAAATCCAGCACTTTGGAGGCCAAAGCGGGTGG 1879
QY 1861 ATGGCTTGAGCCTGAGAGTGAAGAGATGCGGCAAGTCTCAGAGCCCTGTGCGCTC 1920
Db 1880 ATGGCTTGAGCCTGAGAGTGAAGAGATGCGGCAAGTCTCAGAGCCCTGTGCGCTC 1939
QY 1921 GGGGCTCTCTGTGCTGGGCTCCCACTTGGTGGCACTTGAAGAGCCCTTCAAGCCACCG 1980
Db 1940 GGGGCTCTCTGTGCTGGGCTCCCACTTGGTGGCACTTGAAGAGCCCTTCAAGCCACCG 1999
QY 1981 CTGCACTGTGGAGCCCTTCTGTGGCTGGCCAAAGCCCAAGCCGGCTCCCTCAGCTTGC 2040
Db 2000 CTGCACTGTGGAGCCCTTCTGTGGCTGGCCAAAGCCCAAGCCGGCTCCCTCAGCTTGC 2059
QY 2041 AGGAGGTGTGAGGAGGAGGCTCAAGCAGAAACCGGGGCTGGCCAGAGGGGCTTGGCGGC 2100
|||||

Db 2060 AGGAGGTGTGAGGAGGAGAGGCTCAAGCAGGAACCGGGGCTGGCCAGGGGCTTGGCGGC 2119
QY 2101 CAGCTGAGATTCGCGGTGGGCGTGGCTTGGCGGGGCCCGCACTCGGAGAGCGGGCCAG 2160
Db 2120 CAGCTGAGATTCGCGGTGGGCGTGGCTTGGCGGGGCCCGCACTCGGAGAGCGGGCCAG 2179
QY 2161 CCGTTCAGAGGCCCGGGCAATGAGAGGCTTGAACCCGGGGCCAGCGGCTGGGAGGCTGT 2220
Db 2180 CCGTTCAGAGGCCCGGGCAATGAGAGGCTTGAACCCGGGGCCAGCGGCTGGGAGGCTGT 2239
QY 2221 ACTGGGTGCCCCACAGTGCAGGCCCGGGCGGTGTGCTGTGCTGATTTTCTCACTGGGC 2280
Db 2240 ACTGGGTGCCCCACAGTGCAGGCCCGGGCGGTGTGCTGTGCTGATTTTCTCACTGGGC 2299
QY 2281 CTTAGCAGCCTTCCCGGGGAGGAGCTCGGGACCTGCAGCCGCCATGCTGAGCCTTCC 2340
Db 2300 CTTAGCAGCCTTCCCGGGGAGGAGGCTCGGGACCTGCAGCCGCCATGCTGAGCCTTCC 2359
QY 2341 CCTTCATGGGCTCCTGTGCGGCCGAGGCTTCCCGAGAGCAGACACCCCTGTCCACAG 2400
Db 2360 CCTTCATGGGCTCCTGTGCGGCCGAGGCTTCCCGAGAGCAGACACCCCTGTCCACAG 2419
QY 2401 CGCCAGTCCCATGAGCAGCAGAGGCTGAGAGTGGGGGCGGACGGGACCGGAGCTGG 2460
Db 2420 CGCCAGTCCCATGAGCAGCAGAGGCTGAGAGTGGGGGCGGACGGGACCGGAGCTGG 2479
QY 2461 CAGGACGCTACCCCTGCGAGCCCTGTGCGGAATCACTGGGTGAAGCCAGCTGGGCTCT 2520
Db 2480 CAGGACGCTACCCCTGCGAGCCCTGTGCGGAATCACTGGGTGAAGCCAGCTGGGCTCT 2539
QY 2521 GAGTCTGGTGAAGCTTGGAGAACCTTATGCTAGCTCAAGGATCGTAATTAACCAAT 2580
Db 2540 GAGTCTGGTGAAGCTTGGAGAACCTTATGCTAGCTCAAGGATCGTAATTAACCAAT 2599
QY 2581 CAGCACCCTGTGTAGCTCAGGCTGTGTGAATCAGCAATCCACACTGTATATAGCT 2640
Db 2600 CAGCACCCTGTGTAGCTCAGGCTGTGTGAATCAGCAATCCACACTGTATATAGCT 2659
QY 2641 ACTGTATGGGGCTTGGAGAACCTTATGCTAGCTCAAGGATCGTAATTAACCAATC 2700
Db 2660 ACTGTATGGGGCTTGGAGAACCTTATGCTAGCTCAAGGATCGTAATTAACCAATC 2719
QY 2701 GGCACCTGTATCTAGCTCAAGGTTGTAAACACCAATAGCAACCTGTGTAGCTC 2760
Db 2720 GGCACCTGTATCTAGCTCAAGGTTGTAAACACCAATAGCAACCTGTGTAGCTC 2779
QY 2761 AGGTATGTGAATCAGCAATCGACAGCTGTATCTGGCTTCAATGGGCATCCGGT 2820
Db 2780 AGGTATGTGAATCAGCAATCGACAGCTGTATCTGGCTTCAATGGGCATCCGGT 2839
QY 2821 GAAGAGCACCAACAGAGGCTTGTGTGAGCAATTAAGCTTCTATCACTGGGTGACAGT 2880
Db 2840 GAAGAGCACCAACAGAGGCTTGTGTGAGCAATTAAGCTTCTATCACTGGGTGACAGT 2899
QY 2881 GGGCTGAGTCCGAAAGAGAGTCAAGCAAGGAGATTAAGGTTGGGCCGTTTATAGAT 2940
Db 2900 GGGCTGAGTCCGAAAGAGAGTCAAGCAAGGAGATTAAGGTTGGGCCGTTTATAGAT 2959
QY 2941 TTGGGTAGTAAAGAAATTTACAGTCAAAAGGGGCTTGTCTCTGGCGGGCAGAGTGG 3000
Db 2960 TTGGGTAGTAAAGAAATTTACAGTCAAAAGGGGCTTGTCTCTGGCGGGCAGAGTGG 3019
QY 3001 GGGGTGCCAAGTGTCTCAGTGGGGGCTCTTTTATAGCAGAGATGAGCGAAGAAAGAGCT 3060
Db 3020 GGGGTGCCAAGTGTCTCAGTGGGGGCTCTTTTATAGCAGAGATGAGCGAAGAAAGAGCT 3079
QY 3061 TTCAAGATTAATGTATCAATTAAGCAAGAGACCGGCATTTACACCTTTTGTGTG 3120
Db 3080 TTCAAGATTAATGTATCAATTAAGCAAGAGACCGGCATTTACACCTTTTGTGTG 3139
QY 3121 GAATGTATCAGTTAATTTGGGGCAGGCAATTTCACTTCTTTGTGATTTCTCAGTTAC 3180
Db 3140 GAATGTATCAGTTAATTTGGGGCAGGCAATTTCACTTCTTTGTGATTTCTCAGTTAC 3199
|||||

3181 TTCAGGCATCTGGGCGTATATGTCAGTTACAGGGATGCGATGGCTTGGCTGGCT 3240
3200 TTCAGGCATCTGGGCGTATATGTCAGTTACAGGGATGCGATGGCTTGGCTGGCT 3259
3241 CAGAGCTTGCAGCTACTCTGGTGGGCGCTTGAGAAATGTTGTGTGCACCTGTAT 3300
3260 CAGAGCTTGCAGCTACTCTGGTGGGCGCTTGAGAAATGTTGTGTGCACCTGTAT 3319
3301 CTAGTTAATCTAGTGGGACGTGGAAACCTTGTGTCTAGCTCAGCGGATTTGAAACGA 3360
3320 CTAGTTAATCTAGTGGGACGTGGAAACCTTGTGTCTAGCTCAGCGGATTTGAAACGA 3379
3361 CCAATCAGCGCGCTTCAAAAACAGCACTGGGCTCTACCAATTCAGCAGAGATGGGTGG 3420
3380 CCAATCAGCGCGCTTCAAAAACAGCACTGGGCTCTACCAATTCAGCAGAGATGGGTGG 3439
3421 GGCCAGATAGAGATAAAGCAGAGCTGCCGAGCAGCAGTGGCAGCGCAGATCC 3480
3440 GGCCAGATAGAGATAAAGCAGAGCTGCCGAGCAGCAGTGGCAGCGCAGATCC 3499
3481 CTATCCACAATATGGCAGCTTGTCTTGTGTGGATAAATCTTGTACTGTCTGG 3540
3500 CTATCCACAATATGGCAGCTTGTCTTGTGTGGATAAATCTTGTACTGTCTGG 3559
3541 CTTTGTGGTCCACACTGCTTTTATAGCTGTAACTCAACAGAGGTCTGCAGCTTC 3600
3560 CTTTGTGGTCCACACTGCTTTTATAGCTGTAACTCAACAGAGGTCTGCAGCTTC 3619
3601 ACTCTGAAGCCACTAAGACCAAGCCACCGGAGGAGATGAACTCCGCGCGCT 3660
3620 ACTCTGAAGCCACTAAGACCAAGCCACCGGAGGAGATGAACTCCGCGCGCT 3679
3661 GCCTTAAAGCTATTAACCTACCGGAAAGTCTGACGCTTCACTCTCAGCAGGAGA 3720
3680 GCCTTAAAGCTATTAACCTACCGGAAAGTCTGACGCTTCACTCTCAGCAGGAGA 3739
3721 CCAGAACCCCAAGAAAGAACTCGAAACATCTGAAACATCAGAAAGAACT 3780
3740 CCAGAACCCCAAGAAAGAACTCGAAACATCTGAAACATCAGAAAGAACT 3799
3781 CCAGATGCACCACTTAAAGAGCTGTAACTACCTGCGAGGGTCCGCGCTTCTTCTG 3840
3800 CCAGATGCACCACTTAAAGAGCTGTAACTACCTGCGAGGGTCCGCGCTTCTTCTG 3859
3841 AAGTCAGTGAAGCAAGCACTCACCAGTTTGGACACAAGCCAGAGTTTGAATCAGC 3900
3860 AAGTCAGTGAAGCAAGCACTCACCAGTTTGGACACAAGCCAGAGTTTGAATCAGC 3919
3901 CTGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTACAAAAATTTGGCGG 3960
3920 CTGGCAACATGATGAATGCCCTCTCTGCAAAAAAAAATTTACAAAAATTTGGCGG 3979
3961 AGCATGTGTGTCGCTGTGCTGCTCAGCTACCGGGAGGCTAAAGTGGAGATCGCT 4020
3980 AGCATGTGTGTCGCTGTGCTGCTCAGCTACCGGGAGGCTAAAGTGGAGATCGCT 4039
4021 TGAAGCTGGAGGTGAAGACCTGCACTGAGCTGTGATTTGACACAGCCCTCTAGGCTGGG 4080
4040 TGAAGCTGGAGGTGAAGACCTGCACTGAGCTGTGATTTGACACAGCCCTCTAGGCTGGG 4099
4081 GGACAGCTGAGACCTGTTTCCCTCGCAAAAAAAAATTTGACAAAAAGTATTAAGAGT 4140
4100 GGACAGCTGAGACCTGTTTCCCTCGCAAAAAAAAATTTGACAAAAAGTATTAAGAGT 4159
4141 GCCTGATATGCTAGGCGCAGTGGCTCATGCTGTATTCAGCACTTTGGGAAGCCGAG 4200
4160 GCCTGATATGCTAGGCGCAGTGGCTCATGCTGTATTCAGCACTTTGGGAAGCCGAG 4219
4201 GCGGGGGGCTACCTTAAGGTCAAGGAGTGTGAGACACAGCTGGCAACATGGAAGGCC 4260
4220 GCGGGGGGCTACCTTAAGGTCAAGGAGTGTGAGACACAGCTGGCAACATGGAAGGCC 4279

4261 ATCTCTTCAAAAAATTAACAAATTAAGCCGGCTGNGGGGGGCGAGTGGAGCATGCCGTAA 4320
4280 ATCTCTTCAAAAAATTAACAAATTAAGCCGGCTGNGGGGGGCGAGTGGAGCATGCCGTAA 4339
4321 TCCAGCTACTCAGAGAGGCTGAGGCGAGAAATCACTTTGAAACCCAGAGAGCGCGGTTGC 4380
4340 TCCAGCTACTCAGAGAGGCTGAGGCGAGAAATCACTTTGAAACCCAGAGAGCGCGGTTGC 4399
4381 AGTGAAGCAGATCGTCCATTGCACTCCACCCTCCAGCTGGGCAACAAGGCCAA 4440
4400 AGTGAAGCAGATCGTCCATTGCACTCCACCCTCCAGCTGGGCAACAAGGCCAA 4459
4441 CTTCTGCTTAAAAAAAATTAAGAAATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4500
4460 CTTCTGCTTAAAAAAAATTAAGAAATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4519
4501 TTGCGAGGCAACATGTTTAAAGATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4560
4520 TTGCGAGGCAACATGTTTAAAGATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4579
4561 CACCTCAAGGCGCAGGCTGCACTGCTCATGCTATATCCAGCACTTTGGAGCGCAG 4620
4580 CACCTCAAGGCGCAGGCTGCACTGCTCATGCTATATCCAGCACTTTGGAGCGCAG 4639
4621 GCGGGGATCACTGAGGCTCAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4680
4640 GCGGGGATCACTGAGGCTCAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4699
4681 CCCACCTTCAAAAAATTAACAAATTAAGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4740
4700 CCCACCTTCAAAAAATTAACAAATTAAGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4759
4741 CTACTTGGAGGCTGAGGCGAGAAATCACTGAACACAGGAGCGGAGGAGTGTAGTGA 4800
4760 CTACTTGGAGGCTGAGGCGAGAAATCACTGAACACAGGAGCGGAGGAGTGTAGTGA 4819
4801 CCGAGATGCTGATGCTGCACTCAGCTCAGCTGAGAAATGAGGAACTCATCTCAAAAAAAT 4860
4820 CCGAGATGCTGATGCTGCACTCAGCTCAGCTGAGAAATGAGGAACTCATCTCAAAAAAAT 4879
4861 AACAAACAAACCACTCTCTACTCCAGAGGAGTGGGATGAGAGCTGGGCAATCACT 4920
4880 AACAAACAAACCACTCTCTACTCCAGAGGAGTGGGATGAGAGCTGGGCAATCACT 4939
4921 GCAAGGCTGAGGCGACACAGAGTAAAGGCGGAGCTGAGAGACCGGAGCAGATTAACAGT 4980
4940 GCAAGGCTGAGGCGACACAGAGTAAAGGCGGAGCTGAGAGACCGGAGCAGATTAACAGT 4999
4981 TGTGAGATCAGTGTGAGATCAGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040
5000 TGTGAGATCAGTGTGAGATCAGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5059
5041 AGCAGAGATGAGGCGCCCATTCAGATCAGCACTTCATCTCATCCAGAGATGTCTGTT 5100
5060 AGCAGAGATGAGGCGCCCATTCAGATCAGCACTTCATCTCATCCAGAGATGTCTGTT 5119
5101 CTTGGCAGCTGGGGAATTTAGACAGAAAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5160
5120 CTTGGCAGCTGGGGAATTTAGACAGAAAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5179
5161 TGGCCAGGAGGCGCTTGTGGCTGTAGAAAAATTTAGAGGCTTGGAGCCGGGACAGGCTGGC 5220
5180 TGGCCAGGAGGCGCTTGTGGCTGTAGAAAAATTTAGAGGCTTGGAGCCGGGACAGGCTGGC 5239
5221 TCACGCTGTATTCACAGCACTTTGGAGGCGGAGGCGGAGTTCACAGGCTCAGGAGA 5280
5240 TCACGCTGTATTCACAGCACTTTGGAGGCGGAGGCGGAGTTCACAGGCTCAGGAGA 5299
5281 TGTGACCATCTCTGCTTAACAGCGTGAACCCCGCTCTCTACTTAAAAATTAACAAAAATTTGG 5340
5300 TGTGACCATCTCTGCTTAACAGCGTGAACCCCGCTCTCTACTTAAAAATTAACAAAAATTTGG 5359
5341 CCGGGAGTGGTGGGCGACCTGTAGTTCACACTCTGAGGAGGCTGAGGAGGAGTGA 5400

|||||
Db 5360 CCGGGCAGTGTGGGGGCGACCTGTAGTTCAGCTACTCGGAGGCTTAGCAGAGAAATG 5419
QY 5401 GCGTGAACCCGAGAGGAGGAGATTTCAGTACGCGAGATGCGGCACCTGCACCTCACCT 5460
|||||
Db 5420 GCGTGAACCCGAGAGGAGGAGATTTCAGTACGCGAGATGCGGCACCTGCACCTCACCT 5479
QY 5461 GGGCGAGAGAGCAAGACTCCATCTCGAAAAAAGAAAAAGAAAAAGCTTCAGTCTGAGCCAGA 5520
|||||
Db 5480 GGGCGAGAGAGCAAGACTCCATCTCGAAAAAAGAAAAAGAAAAAGCTTCAGTCTGAGCCAGA 5539
QY 5521 GGGCGAGGCTGATTTCTGTACCTTACATGACCTTGGGCAAGGCAATTCCTTCCCTGGC 5580
|||||
Db 5540 GGGCGAGGCTGATTTCTGTACCTTACATGACCTTGGGCAAGGCAATTCCTTCCCTGGC 5599
QY 5581 CCAGTTCACGGGGTGGAAATCGACTCCAGAGTCCCTTCAGACATTAACGCTGCATGGTTTC 5640
|||||
Db 5600 CCAGTTCACGGGGTGGAAATCGACTCCAGAGTCCCTTCAGACATTAACGCTGCATGGTTTC 5659
QY 5641 TAAAGATGAGAAATGGGGCAGTTTCCCTCTCTCACCAGCCCGGTCCATTCAGAGT 5700
|||||
Db 5660 TAAAGATGAGAAATGGGGCAGTTTCCCTCTCTCACCAGCCCGGTCCATTCAGAGT 5719
QY 5701 GAATGACACGAGGAATGACGCTGCTCCAAATCCCGCAGTTCCAAAGCCCTGGGGAGCCGTAC 5760
|||||
Db 5720 GAATGACACGAGGAATGACGCTGCTCCAAATCCCGCAGTTCCAAAGCCCTGGGGAGCCGTAC 5779
QY 5761 TGTACAGGCTGTGCACAGAGAGGTGAAGTCAAGTGAAGCCAAATGCTCGAAGGGCTTTG 5820
|||||
Db 5780 TGTACAGGCTGTGCACAGAGAGGTGAAGTCAAGTGAAGCCAAATGCTCGAAGGGCTTTG 5839
QY 5821 CCTCATTCGGGACAGAGAAATCCGCTTCTCTGCTCTACCGGGATTTCTAGAGGGCTTTAGC 5880
|||||
Db 5840 CCTCATTCGGGACAGAGAAATCCGCTTCTCTGCTCTACCGGGATTTCTAGAGGGCTTTAGC 5899
QY 5881 CGAATGATGATGGGGGGGGGGGGGCTTCTGGGGGAGTTCCCGAGTCAATCAACTTGGGA 5940
|||||
Db 5900 CGAATGATGATGGGGGGGGGGGGGCTTCTGGGGGAGTTCCCGAGTCAATCAACTTGGGA 5959
QY 5941 CAGGACAGCTCGAATCTTTCAGATGCTCTATCCAAATGTGGGGTGGGACAGGACCAA 6000
|||||
Db 5960 CAGGACAGCTCGAATCTTTCAGATGCTCTATCCAAATGTGGGGTGGGACAGGACCAA 6019
QY 6001 GACCCATGCTCTATCTCAAGGTAGGGGCTCAGAGAGTCTCCAGAGAGGAGGCTCCGG 6060
|||||
Db 6020 GACCCATGCTCTATCTCAAGGTAGGGGCTCAGAGAGTCTCCAGAGAGGAGGCTCCGG 6079
QY 6061 AGAGTTTGGGGGTAGGAATGGAGCAACAGGCTTCTTTTCTCTTGAATTTGGG 6120
|||||
Db 6080 AGAGTTTGGGGGTAGGAATGGAGCAACAGGCTTCTTTTCTCTTGAATTTGGG 6139
QY 6121 GCGTTGGGGGAGAGGCTTTCAGATCCAAAGAGAGGGGCAAGAGGACTCCCCACAG 6180
|||||
Db 6140 GCGTTGGGGGAGAGGCTTTCAGATCCAAAGAGAGGGGCAAGAGGACTCCCCACAG 6199
QY 6181 TCTGCCAGAGAGAGAGGAGAGCCCGAGCTCAGCTCCCACTTCCCCACAGGCT 6235
|||||
Db 6200 TCTGCCAGAGAGAGAGGAGAGCCCGAGCTCAGCTCCCACTTCCCCACAGGCT 6254

RESULT 3

US-09-948-128-259
; Sequence 259, Application US/09948128
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; TITLE OF INVENTION: 17, METHODS OF DETECTION, AND USES THEREOF
; FILE REFERENCE: CL001294
; CURRENT APPLICATION NUMBER: US/09/948,128
; NUMBER OF SEQ ID NOS: 465
; SEQ ID NO 259
; LENGTH: 4925599

; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(4925599)
; OTHER INFORMATION: n = A,T,C or G
US-09-948-128-259

Query Match 97.1%; Score 6056; DB 40; Length 4925599;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 6204; Conservative 0; Mismatches 20; Indels 16; Gaps 12;

QY 1 GATCAGCTTGAAGAGACAGTATTCAGACCAAGCTGGGACGATAGGAGAGTGTCTACG 60
Db 1720932 GATCAGCTTGAAGAGACAGTATTCAGACCAAGCTGGGACGATAGGAGAGTGTCTACG 1720991
QY 61 AAAAAATCAAAAAATTTATGGCGGCGCATGTGGGCTCAGCTGTGTAATCCGTAATTGGG 120
Db 1720992 AAAAAATCAAAAAATTTATGGCGGCGCATGTGGGCTCAGCTGTGTAATCCGTAATTGGG 1721051
QY 121 ACATCAAGGCAAGTGGATCAGCTTGAAGTCAAGAGTTCGAGACTAGCTGGCCAAATGT 180
Db 1721052 ACATCAAGGCAAGTGGATCAGCTTGAAGTCAAGAGTTCGAGACTAGCTGGCCAAATGT 1721111
QY 181 GAAACCTTATCTCCACTTAAAAAATCAAAAAATTTAGCCAGGATGTGGGACGCTGTA 240
Db 1721112 GAAACCTTATCTCTACT-AAAAAATCAAAAAATTTAGCCAGGATGTGGGACGCTGTA 1721170
QY 241 ATCCGGCTTACTCAGAGGCTGAGAGGAGAGAAATCAGTTGAACCCAGAGGCGGAGGTTG 300
Db 1721171 ATCCGGCTTACTCAGAGGCTGAGAGGAGAGAAATCAGTTGAACCCAGAGGCGGAGGTTG 1721230
QY 301 CAGTGAAGCTGATCAGACCACTGCACTCCAGCTGGGTGACAGAGCAAGACTATCTC 360
Db 1721231 CAGTGAAGCTGATCAGACCACTGCACTCCAGCTGGGTGACAGAGCAAGACTATCTC 1721290
QY 361 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
Db 1721291 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1721350
QY 421 CTACTCAGAGGCTGAGGCTGGAGAGATCAGTTGAAGCTGGGGACAGTAAAGCTCAGTGA 480
Db 1721351 CTACTCAGAGGCTGAGGCTGGAGAGATCAGTTGAAGCTGGGGACAGTAAAGCTCAGTGA 1721409
QY 481 GCCAAGATCATGCCATCAGCTCCAGCTGGGCAACAGAGAGAGACCTGTCTTAATAA 540
Db 1721410 GCCAAGATCATGCCATCAGCTCCAGCTGGGCAACAGAGAGAGACCTGTCTTAATAA 1721469
QY 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
Db 1721470 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1721526
QY 601 ATGTATATATAGTTTGAACACTCAAGATCAGATAGTCAATTTTATAGGCTGTGGGCGT 660
Db 1721527 ATGTATATATAGTTTGAACACTCAAGATCAGATAGTCAATTTTATAGGCTGTGGGCGT 1721586
QY 661 ATGGTCTGTGCATCAATCACTGTGCTCTCTTCTTTCAGCAAAAGAGCTATTAACAAT 720
Db 1721587 ATGGTCTGTGCATCAATCACTGTGCTCTCTTCTTTCAGCAAAAGAGCTATTAACAAT 1721646
QY 721 ACATACATGAATTTTATAGACATCGAGATTTGAATTTTATATATATATATATATATATAT 780
Db 1721647 ACATACATGAATTTTATAGACATCGAGATTTGAATTTTATATATATATATATATATATAT 1721706
QY 781 AAAAAATATCTTTTAAAAAATTTTCCCTTAACCATTTTAAAAAGTAAAAAGCGGCA--G 837
Db 1721707 AAAAAATATCTTTTAAAAAATTTTCCCTTAACCATTTTAAAAAGTAAAAAGCGGCAAGGCG 1721766
QY 838 CCGCCATCTGTCAAGCTGTAAATTCAGCACTTTGGAGGCTGAGTGGGACATCACTT 897
Db 1721767 CCGCCATGCGCTCAAGCTGTAAATTCAGCACTTTGGAGGCTGAGTGGGACATCACTT 1721826
QY 898 GAGATCAACAGTTTCGAGACCAAGCTGGCCAAACATAGCAAAACCCCATTTCTACTAAAAAT 957

D	1721827	GAGATCAACAGTGTGAGACGACCTGGCCACATAGCAAAACCCATTCTACTPAAAT	1721866	D	1722906	CCACCGCTGCACTGTGGAGCCCTTCTGGGCTGGCCAAAGCCAGAGCCGCTCCCTCA	1722965
Q	958	AAAAAATTTACTTGGGCTATGTGGTGCACACCTGTGATCCAGCTACTTGGAGGCTAG	1017	Q	2035	GCTTCAGAGGAGGTGTGAGAGGAGAGGCTCAAGCAGAAACCGGGCTCGGCAGGGCTT	2094
D	1721887	AAAAAATTTACTTGGGCTATGTGGTGCACACCTGTGATCCAGCTACTTGGAGGCTAG	1721946	D	1722966	GCTTCAGAGGAGGTGTGAGAGGAGAGGCTCAAGCAGAAACCGGGCTCGGCAGGGCTT	1723025
Q	1018	GCAGAGAAATGCTTGAACCTGTGGAGGAGGAGGTTGCTGAGGCCAACATATGCGACTG	1077	Q	2095	GGGGCCAGCTGTGAGTTCCGGGTGGGCTTGGGCTTGGCGGGCCCGCAGCTCGAGACAGC	2154
D	1721947	GCAGAGAAATGCTTGAACCTGTGGAGGAGGAGGTTGCTGAGGCCAACATATGCGACTG	1722006	D	1723026	GGGGCCAGCTGTGAGATTCCGGGTGGGCTTGGGCTTGGCGGGCCCGCAGCTCGAGACAGC	1723085
Q	1078	CACTCAGGCTGTGGTGAAGAGTGAAGTTCGTCAACGAAAAAAGTGTAAAGC	1137	Q	2155	GGCCAGCCCTGCAGAGCCCGGGCAATGAGAGGCTTAGCACCCGGGCCAGCGGCTGCGA	2214
D	1722007	CACTCAGGCTGTGGTGAAGAGTGAAGTTCGTCAACGAAAAAAGTGTAAAGC	1722066	D	1723086	GGCCAGCCCTGCAGAGCCCGGGCAATGAGAGGCTTAGCACCCGGGCCAGCGGCTGCGA	1723145
Q	1138	CATTCTTAATTCAGTGTACATCAGTGTACATCTAGCTGCTGCTGCTGTGAGG	1197	Q	2215	GGGTCTACTGGTGCCCAAGCAGTCCAGCCCGCGGGCTGTGCTCGATTTCTCA	2274
D	1722067	CATTCTTAATTCAGTGTACATCAGTGTACATCTAGCTGCTGCTGCTGTGAGG	1722126	D	1723146	GGGTCTACTGGTGCCCAAGCAGTCCAGCCCGCGGGCTGTGCTCGATTTCTCA	1723205
Q	1198	CATACCTGAGAAAGTAGAGTTGCTGTGCACAGCAATACATTTCCATTAATAGAC	1257	Q	2275	CTGGGCTTAGCAGGCTTCCCGGGGGGAGGGCTGCGGACCTGCAGCCCGCATGCTGCA	2334
D	1722127	CATACCTGAGAAAGTAGAGTTGCTGTGCACAGCAATACATTTCCATTAATAGAC	1722186	D	1723206	CTGGGCTTAGCAGGCTTCCCGGGGGGAGGGCTGCGGACCTGCAGCCCGCATGCTGCA	1723265
Q	1258	ACTACCAAGTGGCCATCCAGAGAGG-TTTTTTTTTACATCTACACTCCCCCAGCAAC	1316	Q	2335	GCCTCCCTCCATGGGCTCTGTGTGGCCCGAGCCTCCCGCAGCAGCAACCCCTGCT	2394
D	1722187	ACTACCAAGTGGCCATCCAGAGAGG-TTTTTTTTTACATCTACACTCCCCCAGCAAC	1722246	D	1723266	GCCTCCCTCCATGGGCTCTGTGTGGCCCGAGCCTCCCGCAGCAGCAACCCCTGCT	1723325
Q	1317	AAATGAGAGTATTCTCCAGATCTCTTTACAAAGATGCTTAAGCCAGTACAGATGAAAC	1376	Q	2395	CCACAGCGCCAGTCCCATCGACACGAGAGGGCTGAGAAATGGGGGCGACGGCACCGG	2454
D	1722247	AAATGAGAGTATTCTCCAGATCTCTTTACAAAGATGCTTAAGCCAGTACAGATGAAAC	1722306	D	1723326	CCACAGCGCCAGTCCCATCGACACGAGAGGGCTGAGAAATGGGGGCGACGGCACCGG	1723385
Q	1377	AGGAAGTGGAGGGGAGAGCTGCCAGCCCTTCTAACCTGGAAGAAATCCGATAGAGCC	1436	Q	2455	GACTGGCAGGCGACGACCCCTGACAGCCCTGGGGGAGAAATCCATGGGGTGAAGCCAGCTGG	2514
D	1722307	AGGAAGTGGAGGGGAGAGCTGCCAGCCCTTCTAACCTGGAAGAAATCCGATAGAGCC	1722366	D	1723386	GACTGGCAGGCGACGACCCCTGACAGCCCTGGGGGAGAAATCCATGGGGTGAAGCCAGCTGG	1723445
Q	1437	TTCTGAGATCTGGAAGATGAATAAAGGGGCTCTGTGAGGCTGCCCTCTGATGATC	1496	Q	2515	GCTCTGAGTGTGTGAGAGCTTGTGAGAACTTTATGTCTAGCTCAGGAGATCTGAATAC	2574
D	1722367	TTCTGAGATCTGGAAGATGAATAAAGGGGCTCTGTGAGGCTGCCCTCTGATGATC	1722426	D	1723446	GCTCTGAGTGTGTGAGAGCTTGTGAGAACTTTATGTCTAGCTCAGGAGATCTGAATAC	1723505
Q	1497	TGTACACTTGAAGGCTCCAGTCCAGTCTCAGCCCCAGTGTCAATGAGGCAATGATATAG	1556	Q	2575	ACCAATTCAGACCCCTGTCTAGCTCAGGAGTGTGAAATGCACCAATTCACACTGTAT	2634
D	1722427	TGTACACTTGAAGGCTCCAGTCCAGTCTCAGCCCCAGTGTCAATGAGGCAATGATATAG	1722485	D	1723506	ACCAATTCAGACCCCTGTCTAGCTCAGGAGTGTGAAATGCACCAATTCACACTGTAT	1723565
Q	1557	CCCTCAGCTCTGTTGGTCTTATCT-CCCCATGTTGGGCTGAAAGTCTGGATTGAGCC	1615	Q	2635	CTAGCTACCTGATGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATACA	2694
D	1722486	CCCTCAGCTCTGTTGGTCTTATCT-CCCCATGTTGGGCTGAAAGTCTGGATTGAGCC	1722545	D	1723566	CTAGCTACCTGATGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATACA	1723625
Q	1616	GTTATTCAGAGTGTACAGCTTCTGACAGGAAAGTATGTCTACAGAAACAGAGGGCT	1675	Q	2695	CCAATTCGCACTCTGTATCTAGCTCAGGAGTGTGAATGAACACACCAATTCAGACCTGTGTC	2754
D	1722546	GTTATTCAGAGTGTACAGCTTCTGACAGGAAAGTATGTCTACAGAAACAGAGGGCT	1722605	D	1723626	CCAATTCGCACTCTGTATCTAGCTCAGGAGTGTGAATGAACACACCAATTCAGACCTGTGTC	1723685
Q	1676	TGGCAAAATGATCTACTGCAAAATCTTACCTGCTCAGCCAGCCAGCTATGTTCTGATCT	1735	Q	2755	TAGCTCAGGAGTATGTAATGCACCAATGACAGTGTATCTGCTACTTTATGAGCAT	2814
D	1722606	TGGCAAAATGATCTACTGCAAAATCTTACCTGCTCAGCCAGCCAGCTATGTTCTGATCT	1722665	D	1723686	TAGCTCAGGAGTATGTAATGCACCAATGACAGTGTATCTGCTACTTTATGAGCAT	1723745
Q	1736	TGAACAAGTTTTTCACTCTCTGAGGCAATCCCTTGGCTCAACAACACAGCTTGGTATGA	1795	Q	2815	CCGTTGGAAGAGACCAACCAAGAGCTTGTGTGAGCAATTAAGTTCTATCACCTGGGT	2874
D	1722666	TGAACAAGTTTTTCACTCTCTGAGGCAATCCCTTGGCTCAACAACACAGCTTGGTATGA	1722725	D	1723746	CCGTTGGAAGAGACCAACCAAGAGCTTGTGTGAGCAATTAAGTTCTATCACCTGGGT	1723805
Q	1796	CAGATGAATGACG-AACTCCCTTACACCTGTAATCCACAGCTTGGAGGCAAGGC	1854	Q	2875	GCAGGTGGGCTGAGTCCGAAAAAGAGAGTCAACGAGAGGAGATTAAGGGTGGGCGTCTTA	2934
D	1722726	CAGATGAATGACG-AACTCCCTTACACCTGTAATCCACAGCTTGGAGGCAAGGC	1722785	D	1723806	GCAGGTGGGCTGAGTCCGAAAAAGAGAGTCAACGAGAGGAGATTAAGGGTGGGCGTCTTA	1723865
Q	1855	GGGGGAGTGGCTTGAAGCTGAGAGGTGACAGCATGCGGGGAGTCTCAACAGCTCGTTC	1914	Q	2935	TAGGATTTGGGTAGGTAAAGGAAATTAACATCAAAAGGGGTTGTCTCTGCGCGGAG	2994
D	1722786	GGGGGAGTGGCTTGAAGCTGAGAGGTGACAGCATGCGGGGAGTCTCAACAGCTCGTTC	1722845	D	1723866	TAGGATTTGGGTAGGTAAAGGAAATTAACATCAAAAGGGGTTGTCTCTGCGCGGAG	1723925
Q	1915	GCTTCGAGGCGCTCTGCTGCGGTCCCACTTCGAGTGGGCACTTGAAGAGCCCTTCAGC	1974	Q	2995	GAGTGGGGGCTCGCAAGGTGCTCAGTGGGGGCTTTTGTGAGCCAGATGAGCCAGGAAA	3054
D	1722846	GCTTCGAGGCGCTCTGCTGCGGTCCCACTTCGAGTGGGCACTTGAAGAGCCCTTCAGC	1722905	D	1723926	GAGTGGGGGCTCGCAAGGTGCTCAGTGGGGGCTTTTGTGAGCCAGATGAGCCAGGAAA	1723985
Q	1975	CCACCGCTGCACTGTGGAGCCCTTCTGGGCTGGCCAAAGCCAGAGCCGCTCCCTCA	2034	Q	3055	AGGACTTTCAACAAGTAATGTATCATTAATTAAGGCAAGGACCCGCATTTACACCTCTTT	3114
D				D	1723986	AGGACTTTCAACAAGTAATGTATCATTAATTAAGGCAAGGACCCGCATTTACACCTCTTT	1724045

OY	3115	TTGCTGGAAATGTCATCAAGTTAAAGTTGGGGGACGGGCAATTTACTCTTCTTGTGATTTCTTC	3174
Db	1724046	GTGGTGGAAATGTCATCAAGTTAAAGTTGGGGGACGGGCAATTTACTCTTCTTGTGATTTCTTC	1724105
OY	3175	AGTTACTTTCAGAGCCATCTGGGCGCTATATATGTCCAAAGTTTCAGGGGATGCGATGGCTGGCT	3234
Db	1724106	AGTTACTTTCAGAGCCATCTGGGCGCTATATATGTCCAAAGTTTCAGGGGATGCGATGGCTGGCT	1724165
OY	3225	TGGGCTCAGAGAGGCTTGACAGCTACTCTGTGTGGGGCCTTGGAAATGTTTGTGTGCACACT	3294
Db	1724166	TGGGCTCAGAGAGGCTTGACAGCTACTCTGTGTGGGGCCTTGGAAATGTTTGTGTGCACACT	1724229
OY	3295	CTGTATCTAGTAAATCTAGTGGGGGACGGAGAAACCTTGTGTACGTCCAGGATTTGTA	3354
Db	1724226	CTGTATCTAGTAAATCTAGTGGGGGACGGAGAAACCTTGTGTACGTCCAGGATTTGTA	1724285
OY	3355	AACGCACCAATCAGGGCCCTGTCAAAAACAGACACTGGCTCTTACCAATCAGCAGATGT	3414
Db	1724286	AACGCACCAATCAGGGCCCTGTCAAAAACAGACACTGGCTCTTACCAATCAGCAGATGT	1724345
OY	3415	GGGTGGGGCCAGATPAGAGAAATPAAAGAGGCTGCCGAGCAGCAGTGGCAAGCGCGAC	3474
Db	1724346	GGGTGGGGCCAGATPAGAGAAATPAAAGAGGCTGCCGAGCAGCAGTGGCAAGCGCGAC	1724405
OY	3475	AGGTCCCTATCCACAAATATGGCAGCTTGTCTTTGCTGTTTGGATTAATCTTGTCTAC	3534
Db	1724406	AGGTCCCTATCCACAAATATGGCAGCTTGTCTTTGCTGTTTGGATTAATCTTGTCTAC	1724465
OY	3535	TGCTCGCTTTTGGGTCCTCCACTGCTTTTATGAGCTGTAAACATCTACCAGCAAGTCTGC	3594
Db	1724466	TGCTCGCTTTTGGGTCCTCCACTGCTTTTATGAGCTGTAAACATCTACCAGCAAGTCTGC	1724525
OY	3595	AGCTTCACTCCTGGAAGCCACTPAGACACGAGCCACCGGGAGATPAAACATCTCCGC	3654
Db	1724526	AGCTTCACTCCTGGAAGCCACTPAGACACGAGCCACCGGGAGATPAAACATCTCCGC	1724585
OY	3655	CGCGTGGCTTAAAGCTATTAACACTTACCGCGAAAGTCTCAGCTTACTCTCAGCCA	3714
Db	1724586	CGCGTGGCTTAAAGCTATTAACACTTACCGCGAAAGTCTCAGCTTACTCTCAGCCA	1724645
OY	3715	GCGAGACCAGCAACCCACAGAGGAAGAAACTCGAACATCTGAAATCAGAAAGAA	3774
Db	1724646	GCGAGACCAGCAACCCACAGAGGAAGAAACTCGAACATCTGAAATCAGAAAGAA	1724705
OY	3775	CAAACTCCAGATGCAACACACTTAAAGCTGTAAACACTCAGGAGGGTCCGGGCTTCC	3834
Db	1724706	CAAACTCCAGATGCAACACACTTAAAGCTGTAAACACTCAGGAGGGTCCGGGCTTCC	1724765
OY	3835	TTCTTGAAGTCAAGTAGAGACCAAGCACTACAGTTTGGAGACCAAGCCAGAGCTTGGAG	3894
Db	1724766	TTCTTGAAGTCAAGTAGAGACCAAGCACTACAGTTTGGAGAGCAAGCCAGAGCTTGGAG	1724825
OY	3895	ATCAGCCCTGGGCAACATGATGAAATGCCCCTCTGTGCAAAAAAAAATTTACAAAAAT	3954
Db	1724826	ATCAGCCCTGGGCAACATGATGAAATGCCCCTCTGTGCAAAAAAAAATTTACAAAAAT	1724885
OY	3955	TGGGGGAGCAATGTGTGTCCGTGTGTGTCCCACTACCGGGAGGCTTAAAGTGGAGG	4014
Db	1724886	TGGGGGAGCAATGTGTGTCCGTGTGTGTCCCACTACCGGGAGGCTTAAAGTGGAGG	1724945
OY	4015	ATCGCTTAGGCTGGGAGGTGAAGACATGACATGAGCTGTGATTTTACACAGCCCTGTAG	4074
Db	1724946	ATCGCTTAGGCTGGGAGGTGAAGACATGACATGAGCTGTGATTTTACACAGCCCTGTAG	1725005
OY	4075	GCTTGGGGGACAGACTGAGACCCCTGTTTCCCTCCGCAAAAAAAATTTGACAAAAGTATATA	4134
Db	1725006	GCTTGGGGGACAGACTGAGACCCCTGTTTCCCTCCGCAAAAAAAATTTGACAAAAGTATATA	1725065
OY	4135	AGAGTGTCTGTATATATGCTTAGGGCCAGTGTGTCTATGCTGTATATCCACAGCACTTTGGGAA	4194
Db	1725066	AGAGTGTCTGTATATATGCTTAGGGCCAGTGTGTCTATGCTGTATATCCACAGCACTTTGGGAA	1725125

QY	4195	UCCGAGGGGGGGTGCACCTTAAGTGCAGAGTGTGAGACCAAGCTTGGCCAACTGGAGA	4254
Db	1725126	GCCGAGGGGGGGGTGCACCTTAAGTGCAGAGTGTGAGACCAAGCTTGGCCAACTGGAGA	1725185
QY	4255	AAGCCATCTCTTCTTAATAAATACAAAATATACCAGCTGTGGGGGACAGTGGTGAAGTC	4314
Db	1725186	AAGCCATCTCTTCTTAATAAATACAAAATATACCAGCTGTGGGGGACAGTGGTGAAGTC	1725245
QY	4315	CTGTAAATCCAGCTACTCAGGAGGCTGAGGAGAGAAATCACTTGAACCCAGAGGGGCC	4374
Db	1725246	CTGTAAATCCAGCTACTCAGGAGGCTGAGGAGAGAAATCACTTGAACCCAGAGGGGCC	1725305
QY	4375	GGTTCGATGAGCCGAGATGCTGGCATTTGCACCTCCACCACTCCAGCTGGGGCAACAAGA	4434
Db	1725306	GGTTCGATGAGCCGAGATGCTGGCATTTGCACCTCCACCACTCCAGCTGGGGCAACAAGA	1725365
QY	4435	GCCAAATCTGTCTTAAAAAAGTGGCCGAGACTATTAAGAGGTGTGCATG	4494
Db	1725366	GCCAAATCTGTCTTAAAAAAGTGGCCGAGACTATTAAGAGGTGTGCATG	1725424
QY	4495	CAATAGTTGCCAGGCACATGTTTAAGAAATGTGGAGCTCCTGGCTTCATGGTCTGTTA	4554
Db	1725425	CAATAGTTGCCAGGCACATGTTTAAGAAATGTGGAGCTCCTGGCTTCATGGTCTGTTA	1725484
QY	4555	AAAACCCCACTCAAGGCCAGGTGCATGGCTCATATATATCCAGCACTTTGGGAG	4614
Db	1725485	AAAACCCCACTCAAGGCCAGGTGCATGGCTCATATATATCCAGCACTTTGGGAG	1725544
QY	4615	GCCGAGGGGGGTGATCACTGAGGTGAGAGTTCAGAACCAAGCTGACCAACAATGG	4674
Db	1725545	GCCGAGGGGGGTGATCACTGAGGTGAGAGTTCAGAACCAAGCTGACCAACAATGG	1725604
QY	4675	TGAAATCCCACTCTACTAAAAAATACAAAATTAAGATGAGCATGGTGTGATGCTGTAA	4734
Db	1725605	TGAAATCCCACTCTACTAAAAAATACAAAATTAAGATGAGCATGGTGTGATGCTGTAA	1725664
QY	4735	TCCCACTACTTTGGAGGCTGAGGCAGGAAATCACTGAACCAAGAGCGGAGGTTGT	4794
Db	1725665	TCCCACTACTTTGGAGGCTGAGGCAGGAAATCACTGAACCAAGAGCGGAGGTTGT	1725724
QY	4795	AGTGAGCCGAGATGCTGGCATTTGTCACCTGCAGCAATGAGACGAAATCTCAATCTCA	4854
Db	1725725	AGTGAGCCGAGATGCTGGCATTTGTCACCTGCAGCAATGAGACGAAATCTCAATCTCA	1725784
QY	4855	AAAAACAACAACCAACCACTCTACTCCAGGAGGCTGGGTACAGACTGGGCCAC	4914
Db	1725785	AAAAACAACAACCAACCACTCTACTCCAGGAGGCTGGGTACAGACTGGGCCAC	1725844
QY	4915	ATCAGTGCAGAGTCTGTGAGCCACAGAGCTAAGGGGGAGCTCAGAGACCGGGACCAAGTA	4974
Db	1725845	ATCAGTGCAGAGTCTGTGAGCCACAGAGCTAAGGGGGAGCTCAGAGACCGGGACCAAGTA	1725904
QY	4975	ACAGTGTGTGAGATCAGTGTGTGAGATCAGAGCTGCCATTTGGTGAACCAAGGGGG	5034
Db	1725905	ACAGTGTGTGAGATCAGTGTGTGAGATCAGAGCTGCCATTTGGTGAACCAAGGGGG	1725966
QY	5035	CCCCAAGCACAGAGATGGCCCCATCCAGTCACACATCCACTTTCATCCAGAGATGT	5094
Db	1725965	CCCCAAGCACAGAGATGGCCCCATCCAGTCACACATCCACTTTCATCCAGAGATGT	1726024
QY	5095	CTGTTTCTTGGCACGCTGGGGTAAATTAAGACAGAGGTGACACTTTGGGTGTGGTCAG	5154
Db	1726025	CTGTTTCTTGGCACGCTGGGGTAAATTAAGACAGAGGTGACACTTTGGGTGTGGTCAG	1726084
QY	5155	TCACACATGCCCCAGGACAGGCTTTGTGGCTTAAAGAAACGTTCAAGGCTTAGGCCGGCAC	5214
Db	1726085	TCACACATGCCCCAGGACAGGCTTTGTGGCTTAAAGAAACGTTCAAGGCTTAGGCCGGCAC	1726144
QY	5215	GGTGAGCTACGCTGTATATCCACACTTTTGGAGGGCCAGGGGGGTGAATCAGAGATGC	5274
Db	1726145	GGTGAGCTACGCTGTATATCCACACTTTTGGAGGGCCAGGGGGGTGAATCAGAGATGC	1726204
QY	5275	AGGAGATCTGATCCATCCTGGCTTAACAGAGGTGAACCCCGCTTCTACTAAATAATACAAA	5334

QY	898	GAGATCAACAGTTGGCAGACACAGCCTGGCCAAACATAGCAAAAACCCATTTCCTACTAAAT	957
Db	3470	GAGATCAACAGTTTCAGACACAGCCGGCCAACTAGCAAAAACCCATTTCCTACTAAAT	3529
QY	958	AAAAAATTAGCTGGGCATAGGGGACACCTGTGATCCAGACTTTGGGAGCTAG	1017
Db	3530	AAAAAATTAGCTGGGCATAGGGGACACCTGTGATCCAGACTTTGGGAGCTAG	3589
QY	1018	GCAGAGAAATGCTTGAACCTTGGGAACGGAGGTGACGTAGAGCCACATCATGACCTG	1077
Db	3590	GCAGAGAAATGCTTGAACCTTGGGAACGGAGGTGACGTAGAGCCACATCATGACCTG	3649
QY	1078	CACATCCAGCCGGGGACAGAGATGAGACTTGTCTCAACGAAAAAAGTGTAAAGC	1137
Db	3650	CACATCCAGCCGGGGACAGAGATGAGACTTGTCTCAACGAAAAAAGTGTAAAGC	3709
QY	1138	CATTCTTAATTCAGTACATACAGTACATACATACATAGCTTCGCGTACTCGCTGTAG	1197
Db	3710	CATTCTTAATTCAGTACATACAGTACATACATAGCTTCGCGTACTCGCTGTAG	3765
QY	1198	CATACCTAGAAATGAGATGCTGCTGACAGGACATACACATTTCCACATTAATAGAC	1257
Db	3770	CATACCTAGAAATGAGATGCTGCTGACAGGACATACACATTTCCACATTAATAGAC	3829
QY	1258	ACTACCAAGTTGGCCATCCAGAGAG - TTTTTTTTTTAAATCTACACCCCCACAGAC	1316
Db	3830	ACTACCAAGTTGGCCATCCAGAGAGTTTTTTTTTTAAATCTACACCCCCACAGAC	3889
QY	1317	AAATGAGAGTTACTCCAGATCCTTTTACAAAGATGCTTAAGCCCGTACAGATGAAC	1376
Db	3890	AAATGAGAGTTACTCCAGATCCTTTTACAAAGATGCTTAAGCCCGTACAGATGAAC	3949
QY	1377	AGGAAGTGGGAGGGGAACCTGCGACCCCTCTTAACCATGAAGAAATCCTGTGAGCC	1436
Db	3950	AGGAAGTGGGAGGGGAACCTGCGACCCCTCTTAACCATGAAGAAATCCTGTGAGCC	4009
QY	1437	TTCTGAGATGCGAAGGATGAATAACGGGGGCTCTGTGAGGCTGGCCCTGTACAGATAC	1496
Db	4010	TTCTGAGATGCGAAGGATGAATAACGGGGGCTCTGTGAGGCTGGCCCTGTACAGATAC	4065
QY	1497	TGTGACTTCTGAGCCTCCAGTCCAGCTCTCAGCCCATGTGTATGATGATGATGAG	1556
Db	4070	TGTGACTTCTGAGCCTCCAGTCCAGCTCTCAGCCCATGTGTATGATGATGAG	4128
QY	1557	CCCTACTCTCTGTTTGGCTTTTATTC - CCCCAGTGGGGCTGAAGTCTGGATTTGACC	1615
Db	4129	CCCTACTCTCTGTTTGGCTTTTATTC - CCCCAGTGGGGCTGAAGTCTGGATTTGACC	4188
QY	1616	GTTATTTCAAGATGTACAGCTTTCTTGACAGGAAAGTATGTACAGAAACAGAGGGCT	1675
Db	4189	GTTATTTCAAGATGTACAGCTTTCTTGACAGGAAAGTATGTACAGAAACAGAGGGCT	4248
QY	1676	TGGCAAGATGATCTTAACGTCAAAATCTACCTGGCTCAGCCACAGACTAAGTTCTGATCT	1735
Db	4249	TGGCAAGATGATCTTAACGTCAAAATCTACCTGGCTCAGCCACAGACTAAGTTCTGATCT	4308
QY	1736	TGAACAATTTTTTCACTTCTCTGAGGCAATCCCTGGCTCAACAACACAGAGTTGGTGA	1795
Db	4309	TGAACAATTTTTTCACTTCTCTGAGGCAATCCCTGGCTCAACAACACAGAGTTGGTGA	4368
QY	1796	CAGATGAAATGAGC - AAGTCCCTTACACCTGTATCCACAGACTTTTGGAGGCCAAGC	1854
Db	4369	CAGATGAAATGAGC - AAGTCCCTTACACCTGTATCCACAGACTTTTGGAGGCCAAGC	4428
QY	1855	GGGTGAGATGCTTGAAGCTTGAGAGGTGACAGCATGCGGGCAGTCTCTACAGCCCTGTT	1914
Db	4429	GGGTGAGATGCTTGAAGCTTGAGAGGTGACAGCATGCGGGCAGTCTCTACAGCCCTGTT	4488
QY	1915	GCTCTCGAGGCTCTCTGCTCTGGGCTCCACTTCGCGAGCACTTGAGAGACCTTTAGC	1974
Db	4489	GCTCTCGAGGCTCTCTGCTCTGGGCTCCACTTCGCGAGCACTTGAGAGACCTTTAGC	4548

QY	1975	CCACCGCTGCACTGTGGAGGCCCTTTCTGGGCTGGCCAAAGGCCAGACCGGCTCCCTCA	203
Db	4549	CCACCGCTGCACTGTGGAGGCCCTTTCTGGGCTGGCCAAAGGCCAGACCGGCTCCCTCA	4608
QY	2035	GCTTCAGAGGAGTGTGAGAGGAGAGGCTCAAGCAGAAACGGGGCTCGCACGGCGCTT	2094
Db	4609	GCTTCAGAGGAGTGTGAGAGGAGAGGCTCAAGCAGAAACCGGGCTCGCGCGGCGCTT	4666
QY	2095	GCGGGCCAGCTGGAATTCCGGGTGGGCTGGGCTTGGCGGGCCCGCACCTCGAGCACG	2154
Db	4669	GCGGGCCAGCTGGAATTCCGGGTGGGCTGGGCTTGGCGGGCCCGCACCTCGAGCACG	4726
QY	2155	GGCCAGCCCTGGCAGGCCCCGGGGCAATGAGAGCTTACACCCGGGGCCAGGGGGTGGGA	2214
Db	4729	GGCCAGCCCTGGCAGGCCCCGGGGCAATGAGAGCTTACACCCGGGGCCAGGGGGTGGGA	4786
QY	2215	GGGTGTACTGTGGGTGCCCAAGCAGTGCACGCCCGGGCGGTGTGCTCGCTCGATTTCTCA	2274
Db	4789	GGGTGTACTGTGGGTGCCCAAGCAGTGCACGCCCGGGCGGTGTGCTCGCTCGATTTCTCA	4844
QY	2275	CTGGGCGCTTAGAGCCTTCCCGGGGGGAGGGCTCGGGACCTGCAGCGCCCGCATGCCCTGA	2333
Db	4849	CTGGGCGCTTAGAGCCTTCCCGGGGGGAGGGCTCGGGACCTGCAGCGCCCGCATGCCCTGA	4908
QY	2335	GCCTTCCCTCCATGGGCTCTGTGTGGGCGCCGAGCCTCCCGACAGCAGCACCCCTGCT	2394
Db	4909	GCCTTCCCTCCATGGGCTCTGTGTGGGCGCCGAGCCTCCCGACAGCAGCACCCCTGCT	4966
QY	2395	CCACAGCCCGCCAGTCCCATTCGACACCGCAAGGGCTGAGAAATGGGGGGCGCACGGCACGG	2455
Db	4969	CCACAGCCCGCCAGTCCCATTCGACACCGCAAGGGCTGAGAAATGGGGGGCGCACGGCACGG	5028
QY	2455	GACTGGCAGGAGCTACCCCTGTCAGACCCCTGGTGGCGAATCCACTGGGTGAAGACGACGTGG	2514
Db	5029	GACTGGCAGGAGCTACCCCTGTCAGACCCCTGGTGGCGAATCCACTGGGTGAAGACGACGTGG	5088
QY	2515	GCTCTGAGTCTGGTGGAGACTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATAC	2574
Db	5089	GCTCTGAGTCTGGTGGAGACTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATAC	5148
QY	2575	ACCAATCAGACACCCTGTCTAGCTCAGGAGCTGTGGATGGAACCAATCGACACTGTGAT	2634
Db	5149	ACCAATCAGACACCCTGTCTAGCTCAGGAGCTGTGGATGGAACCAATCGACACTGTGAT	5208
QY	2635	CTAGTACTCTGATGGGGCCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATACA	2694
Db	5209	CTAGTACTCTGATGGGGCCTTGGAGAACTTTATGTCTAGCTCAGGGATGTAAATACA	5268
QY	2695	CCAAATCGGCACCTGTATCTAGCTCAAGGTTGTAAACCAACCAATCGACACCCTGTGTC	2754
Db	5269	CCAAATCGGCACCTGTATCTAGCTCAAGGTTGTAAACCAACCAATCGACACCCTGTGTC	5328
QY	2755	TAGCTCAGGATGTGAAATGCAACCAATGCACAGTCTGTATGTGGCTACTTTCATGGGCAT	2814
Db	5329	TAGCTCAGGATGTGAAATGCAACCAATGCACAGTCTGTATGTGGCTACTTTCATGGGCAT	5388
QY	2815	CCGTGTGAAGAGACACCAACACAGGCTTGTGTGAGCATTTAAAGCTTTATCACCTGGGT	2874
Db	5389	CCGTGTGAAGAGACACCAACACAGGCTTGTGTGAGCATTTAAAGCTTTATCACCTGGGT	5448
QY	2875	GCAGGTGGGCTGAGTCCGAAAAGAGAGTCAACGAGAGGAGATTAAGGGTGGGGCCGTTTAA	2934
Db	5449	GCAGGTGGGCTGAGTCCGAAAAGAGAGTCAACGAGAGGAGATTAAGGGTGGGGCCGTTTAA	5508
QY	2935	TAGATTTTGGGTAGGTAAAGGAAATTTACAGTCAAAAGGGGCTTGTCTCTGCGGGCAG	2994
Db	5509	TAGATTTTGGGTAGGTAAAGGAAATTTACAGTCAAAAGGGGCTTGTCTCTGCGGGCAG	5568
QY	2995	GAGTGGGGGGTGCAGAAAGTGTCTCAGTGGGGGTGCTTTTGTAGCCAGAGATGAGCCAGGAAA	3054
Db	5569	GAGTGGGGGGTGCAGAAAGTGTCTCAGTGGGGGTGCTTTTGTAGCCAGAGATGAGCCAGGAAA	5628
QY	3055	AGGACTTTCACAAGGTATGTCAATTAAGGCAAGGACCCGGCATTTACACTCTTTT	3114

|||||
Db 5629 AGGACTTTACAAAGTATGCATCAATTAAGCCAAAGACCAGCATTTTACACCTCTTTT 5688
OY 3115 GTGGTGAATGTATCATGTTAAGTTGGGGAGGGCATTTACATCTCTTTGTGATTTCTC 3174
Db 5689 GTGGTGAATGTATCATGTTAAGTTGGGGAGGGCATTTTACATCTCTTTGTGATTTCTC 5748
OY 3175 AGTTACTTCAGAGCCATCTGGGCGCTATATGTCAGATTACAGGGGATGCGATGGCT 3234
Db 5749 AGTTACTTCAGAGCCATCTGGGCGCTATATGTCAGATTACAGGGGATGCGATGGCT 5808
OY 3235 TGGGCTCAGAGGCTTGACAGCTACTCTGTGGGGCGCTTGAGAAATGTTTGTGACACT 3294
Db 5809 TGGGCTCAGAGGCTTGACAGCTACTCTGTGGGGCGCTTGAGAAATGTTTGTGACACT 5868
OY 3295 CTGTATCTAGTTATCTAGTGGGAGCGTGGAGAACTTTGTGTACTCTAGGGATTTGTA 3354
Db 5869 CTGTATCTAGTTATCTAGTGGGAGCGTGGAGAACTTTGTGTACTCTAGGGATTTGTA 5928
OY 3355 AACGCACCAATCAGCGCCCTGTCTAAACAGACACTGGGCTCTACCAATCAGAGATGT 3414
Db 5929 AACGCACCAATCAGCGCCCTGTCTAAACAGACACTGGGCTCTACCAATCAGAGATGT 5988
OY 3415 GGGTGGGGCAGATTAAGATAAAGCAGGCTGCCAGCAGCAGAGTGGCAACGCGCAC 3474
Db 5989 GGGTGGGGCAGATTAAGATAAAGCAGGCTGCCAGCAGCAGAGTGGCAACGCGCAC 6048
OY 3475 AGGTCCCTATCCACAAATATGGCAGCTTGTCTTTCTGTCTGCTTAATCTTGTCTAC 3534
Db 6049 AGGTCCCTATCCACAAATATGGCAGCTTGTCTTTCTGTCTGCTTAATCTTGTCTAC 6108
OY 3535 TGGTCCGTTTTTGGGTCACACACTGCTTTATAGCTGTAAACACTCACACAGAAAGTCTGC 3594
Db 6109 TGGTCCGTTTTTGGGTCACACACTGCTTTATAGCTGTAAACACTCACACAGAAAGTCTGC 6168
OY 3595 AGCTTCACTCTCTGAAGCCACTTAAGACACAGCCACCGGAGGAAATGAACAACTCCGGC 3654
Db 6169 AGCTTCACTCTCTGAAGCCACTTAAGACACAGCCACCGGAGGAAATGAACAACTCCGGC 6228
OY 3655 CGGGCTCCCTTAAGAGCTATTAACACTCAGCGGAAAGTCTGACCTTCACTCTCAGCCA 3714
Db 6229 CGGGCTCCCTTAAGAGCTATTAACACTCAGCGGAAAGTCTGACCTTCACTCTCAGCCA 6288
OY 3715 GCGAGACCAAGAACCCACGAAGGAGAAACTGCAACATCTGAACATCAGAAAGAA 3774
Db 6289 GCGAGACCAAGAACCCACGAAGGAGAAACTGCAACATCTGAACATCAGAAAGAA 6348
OY 3775 CAAACTCCAGATGCACCACTTAAGAGCTGTAAACACTCAGTGCAGGGTCCGCGCTTCC 3834
Db 6349 CAAACTCCAGATGCACCACTTAAGAGCTGTAAACACTCAGTGCAGGGTCCGCGCTTCC 6408
OY 3835 TTCTTGAAGTCAGTAGAGACCAAGCACTCACCAAGTTTGGACACAAAGCCAGAGTTTGA 3894
Db 6409 TTCTTGAAGTCAGTAGAGACCAAGCACTCACCAAGTTTGGACACAAAGCCAGAGTTTGA 6468
OY 3895 ATCAGCTCGGGCAACATGATGAATGCCCTCTCGCAAAAAAATTTTACAAAAT 3954
Db 6469 ATCAGCTCGGGCAACATGATGAATGCCCTCTCGCAAAAAAATTTTACAAAAT 6528
OY 3955 TGGCGAGCATGTGTGTCCTGCTGTGTCCAGTACGCGGAGGCTTAAAGTGGAGG 4014
Db 6529 TGGCGAGCATGTGTGTCCTGCTGTGTCCAGTACGCGGAGGCTTAAAGTGGAGG 6588
OY 4015 ATGCTTGAAGCTCTGGAGGTGAAGACTGCAAGTGTATGTTTACACAGCCCTCTAG 4074
Db 6589 ATGCTTGAAGCTCTGGAGGTGAAGACTGCAAGTGTATGTTTACACAGCCCTCTAG 6648
OY 4075 GCTGGGGAGCAGACTGAGACCTGTTTCCCTCGCAAAAAAATTTGCAAAAGTATATA 4134
Db 6649 GCTGGGGAGCAGACTGAGACCTGTTTCCCTCGCAAAAAAATTTGCAAAAGTATATA 6708
OY 4135 AGAGGTGCTGATATGGCTAGGGCAGTGGCTCATGCTGTAACTCCAGCACTTTGGAA 4194
|||||

Db 6709 AGAGGTGCTGATATGGCTAGGCGCAGTGGCTCATGCTGTAACTCCAGCACTTTGGAA 6768
OY 4195 GCCAGAGGGGGGTACCTAAGGTACAGAGTGTGAGACAGCCTGGCCAAAGTGGAGA 4254
Db 6769 GCCAGAGGGGGGTACCTAAGGTACAGAGTGTGAGACAGCCTGGCCAAAGTGGAGA 6828
OY 4255 AAGCCATCTCTTCTTAAATAATACAAATATACCCGCTGTGGGGCAGTGTGAGCAATC 4314
Db 6829 AAGCCATCTCTTCTTAAATAATACAAATATACCCGCTGTGGGGCAGTGTGAGCAATC 6888
OY 4315 CTGTATCCCAAGCTACTCAGAGGCTGAGGCAAGAAATCACTTGAACCCAGAGGCGGC 4374
Db 6889 CTGTATCCCAAGCTACTCAGAGGCTGAGGCAAGAAATCACTTGAACCCAGAGGCGGC 6948
OY 4375 GGTGCACTGAGCGAGATGCGCATGTCACCTCCACCATCCAGCCTGGGCAACAGA 4434
Db 6949 GGTGCACTGAGCGAGATGCGCATGTCACCTCCACCATCCAGCCTGGGCAACAGA 7008
OY 4435 GCCAAACTCTCTTAAATAAATAAAGTCCCTGACATATTAAGGTGTGCAATG 4494
Db 7009 GCCAAACTCTCTT -AAAAAATAAATAAAGTCCCTGACATATTAAGGTGTGCAATG 7067
OY 4495 CAATAGTTGCCAGGCAACATGTTTAAGATGTGAAGTCTGCTTCCATGGTCTGTTA 4554
Db 7068 CAATAGTTGCCAGGCAACATGTTTAAGATGTGAAGTCTGCTTCCATGGTCTGTTA 7127
OY 4555 AAAACCCACCTCAAGGCGAGGTGAGTGCATGCTTATATATCCAGCACTTTGGAG 4614
Db 7128 AAAACCCACCTCAAGGCGAGGTGAGTGCATGCTTATATATCCAGCACTTTGGAG 7187
OY 4615 GCCGAGGGGGGTGATCACTGAGGTGAGAGTTGAGACAGCCTGACCAACCAATGG 4674
Db 7188 GCCGAGGGGGGTGATCACTGAGGTGAGAGTTGAGAGTTGAGACAGCCTGACCAACCAATGG 7247
OY 4675 TGAATCCACCTCTACTTAAATAATACAAATTTAGTGAAGTGTGTGATGCTGTAA 4734
Db 7248 TGAATCCACCTCTACTTAAATAATACAAATTTAGTGAAGTGTGTGATGCTGTAA 7307
OY 4735 TCCCACTACTTGGGAGGCTGAGGCAAGAAATCACTTGAACAGGAGGGGAGGTGT 4794
Db 7308 TCCCACTACTTGGGAGGCTGAGGCAAGAAATCACTTGAACAGGAGGGGAGGTGT 7367
OY 4795 AGTGAGCGAGATGTGTCATTTGACACTCAGCTGAGCAATGAGGAAACTGCATCTCA 4854
Db 7368 AGTGAGCGAGATGTGTCATTTGACACTCAGCTGAGCAATGAGGAAACTGCATCTCA 7427
OY 4855 AAAAAACAACAAACCAACCTCTACTCTCCAGGAGAGTGGGTACAGAGCTGGGCCAC 4914
Db 7428 AAAAAACAACAAACCAACCTCTACTCTCCAGGAGAGTGGGTACAGAGCTGGGCCAC 7487
OY 4915 ATCAGTGCAGGTGCTGAGCCACAGAGCTTAAGCGGAGCTGCAGAGCCGGGACCATTA 4974
Db 7488 ATCAGTGCAGGTGCTGAGCCACAGAGCTTAAGCGGAGCTGCAGAGCCGGGACCATTA 7547
OY 4975 ACAGTGTGAGTACAGTGTGAGTACAGACGTCCTCCATGTGTACCAACAGGGGG 5034
Db 7548 ACAGTGTGAGTACAGTGTGAGTACAGACGTCCTCCATGTGTACCAACAGGGGG 7607
OY 5035 CCCCAGACACAGAGATGGCCCATCCAGTACACACATCCACTTCTCATCAGAGATG 5094
Db 7608 CCCCAGACACAGAGATGGCCCATCCAGTACACACATCCACTTCTCATCAGAGATG 7667
OY 5095 CTGTTTCTTGGCAGCTGGGGTAAATTAAGACAGAGGTGACAGTCTTGGGTGTGTCAG 5154
Db 7668 CTGTTTCTTGGCAGCTGGGGTAAATTAAGACAGAGGTGACAGTCTTGGGTGTGTCAG 7727
OY 5155 TCAGACGTCCCAAGGAGGCTTGTGCTGTATAGAAAACGTTCAAGGCTTAGCGGGGAC 5214
Db 7728 TCAGACGTCCCAAGGAGGCTTGTGCTGTATAGAAAACGTTCAAGGCTTAGCGGGGAC 7787
OY 5215 GGTGCTCAGCGCTGTATATCCACACTTTTGGAGGCGGAGCGGAGTACAGAGGTC 5274
Db 7788 GGTGCTCAGCGCTGTATATCCACACTTTTGGAGGCGGAGCGGAGTACAGAGGTC 7847

[illegible]

OY 3612 CACTAAGACACGAGCCACCAGGAGGAGTAATGAACTCCGGCCGCTGCTTAAAGC 3671
|||||
Db 3057 CACTAAGACACGAGCCACCAGGAGGATTAACACTCCGGCCGCTGCTTAAAGC 3116
|||||
OY 3672 TATTAACCTACCCCGAAGCTCTGACGCTTCACTCTGACGACGAGACCAACCA 3731
|||||
Db 3117 TATTAACCTACCCCGAAGGCTCTGACCTTCACTCTGACGACGAGACCAACCA 3176
|||||
OY 3732 CCAAGAGGAGAAACCTGGAACATCTGAACATCAAGAAACAACTCCAGATGACC 3791
|||||
Db 3177 CCAAGAGGAGAAACCTGGAACATCTGAACATCAAGAAACAACTCCAGATGACC 3236
|||||
OY 3792 ACCTTAAGAGCTGTAACTACTGCGAGGGTCCGGCTCTCTTGAAGTAGAG 3851
|||||
Db 3237 ACCTTAAGAGCTGTAACTACTGCGAGGGTCCGGCTCTCTTGAAGTAGAG 3296
|||||
OY 3852 ACCAAGCACTCACCAAGTTTGGACACAAGCCGAGAGTTTGAGATCAGCCTGGCAACAT 3911
|||||
Db 3297 ACCAAGCACTCACCAAGTTTGGACACAAGCCGAGAGTTTGAGATCAGCCTGGCAACAT 3356
|||||
OY 3912 GATGAATGCCCTCTCTGCAAAAAAATTAACAAAAATTTGGCGGAGCATGTGCT 3971
|||||
Db 3357 GATGAATGCCCTCTCTCTGCAAAAAAATTAACAAAAATTTGGCGGAGCATGTGCT 3416
|||||
OY 3972 CCGTGCCTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGAGATCGCTTGAGCCCTGGGA 4031
|||||
Db 3417 CCGTGCCTGTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGAGATCGCTTGAGCCCTGGGA 3476
|||||
OY 4032 GGTGAAGACTGCACTGAGCTGTGATGTACCAAGCCCTCTAGGCTGGGGAGCACTGA 4091
|||||
Db 3477 GGTGAAGACTGCACTGAGCTGTGATGTACCAAGCCCTCTAGGCTGGGGAGCACTGA 3536
|||||
OY 4092 GACCCTGTTTCCCTCCGCAAAAAATTAACAAAAATTTGAAGAGTGCCTGATATGG 4151
|||||
Db 3537 GACCCTGTTTCCCTCCGCAAAAAATTAACAAAAATTTGAAGAGTGCCTGATATGG 3596
|||||
OY 4152 CTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCCGAGCGGGGCTC 4211
|||||
Db 3597 CTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCCGAGCGGGGCTC 3656
|||||
OY 4212 ACCTAAGGTCAAGAGTGTGAGACCAAGCTTGGCCCAACATGGAAGAGCCCATCTCTTAA 4271
|||||
Db 3657 ACCTAAGGTCAAGAGTGTGAGACCAAGCTTGGCCCAACATGGAAGAGCCCATCTCTTAA 3716
|||||
OY 4272 AAATACAAATTAAGCCGCTGTGGGGAGTGGTGGAGCATGCTGTAATCCAGCTACT 4331
|||||
Db 3717 AAATACAAATTAAGCCGCTGTGGGGAGTGGTGGAGCATGCTGTAATCCAGCTACT 3776
|||||
OY 4332 CAGGAGGCTGAGGCAAGATCATCTTGAACCCAGAGGCGGCTTGCACTGAGCCGAG 4391
|||||
Db 3777 CAGGAGGCTGAGGCAAGATCATCTTGAACCCAGAGGCGGCTTGCACTGAGCCGAG 3836
|||||
OY 4392 ATCGTSCATWTGACATCCACCCTCAGCCTGGGGCAACAAGAGCCAACTCTGTCTTAA 4451
|||||
Db 3837 ATCGTSCATWTGACATCCACCCTCAGCCTGGGGCAACAAGAGCCAACTCTGTCTTAA 3895
|||||
OY 4452 AAAAAAAAAAAGTGCCTGACATATAAGAGTGTGCAATGCAATAGTTGGCAGAGCAA 4511
|||||
Db 3896 AAAAAAAAAAAGTGCCTGACATATAAGAGTGTGCAATGCAATAGTTGGCAGAGCAA 3955
|||||
OY 4512 CATGTTTAAAGATGTGAGAGCTCCTGCTTCATGTGCTGTAAAAAACCCACCTCAAG 4571
|||||
Db 3956 CATGTTTAAAGATGTGAGAGCTCCTGCTTCATGTGCTGTAAAAAACCCACCTCAAG 4015
|||||
OY 4572 CCAAGGAGAGTGGGCTCATGCTTAATATCCACACATTTGGAGGCCGAGGGGTGATC 4631
|||||
Db 4016 CCAAGGAGAGTGGGCTCATGCTTAATATCCACACATTTGGAGGCCGAGGGGTGATC 4075
|||||
OY 4632 ACCTGAGGTCAAGAGTTGAGAGCAAGCTGACACCAACATGTGAAATCCCAACCTTAC 4691
|||||
Db 4076 ACCTGAGGTCAAGAGTTGAGAGCAAGCTGACACCAACATGTGTGAATCCCAACCTTAC 4135
|||||

OY 4692 TAAAAATACAAATTAAGATGAGCATGTGTGTCATGCTGTAAATCCCACTTACTTGGAG 4751
|||||
Db 4136 TAAAAATACAAATTAAGATGAGCATGTGTGTCATGCTGTAAATCCCACTTACTTGGAG 4195
|||||
OY 4752 GCTGAGGAGGAAAAATCACTAGAACCAAGGAGGCGGAGTGTGTAGTGGCAGATGCTG 4811
|||||
Db 4196 GCTGAGGAGGAAAAATCACTAGAACCAAGGAGGCGGAGTGTGTAGTGGCAGATGCTG 4255
|||||
OY 4812 CCAATTGCATCTCAGCCTGAGGAATGAGGAAACTTCATCTCAAAAAAACAACAACAAAA 4871
|||||
Db 4356 CCAATTGCATCTCAGCCTGAGGAATGAGGAAACTTCATCTCAAAAAAACAACAACAAAA 4315
|||||
OY 4872 CCCACTCTCTACCTCCAGAGGAGCTGGGTACAGAGCTGGGCCATCACTGCAAGGTGCTG 4931
|||||
Db 4316 CCCACTCTCTACCTCCAGAGGAGCTGGGTACAGAGCTGGGCCATCACTGCAAGGTGCTG 4375
|||||
OY 4932 AGCCACAGAGCTAAGGCGGAGCTCAGAGACCGCGGACAGATTAACAGTGTGATGATCAG 4991
|||||
Db 4376 AGCCACAGAGCTAAGGCGGAGCTCAGAGACCGCGGACAGATTAACAGTGTGATGATCAG 4435
|||||
OY 4992 TGTGTAGATCAGACGTCCCTGACCATTTGTGACCAAGGAGGCCCCCAAGCACAGAGA 5051
|||||
Db 4436 TGTGTAGATCAGACGTCCCTGACCATTTGTGACCAAGGAGGCCCCCAAGCACAGAGA 4495
|||||
OY 5052 TGGCCCATCCAGTCACACATCCACTTCTCATCAAGATGTCTGTTTCTTGGCACGCT 5111
|||||
Db 4496 TGGCCCATCCAGTCACACATCCACTTCTCATCAAGATGTCTGTTTCTTGGCACGCT 4555
|||||
OY 5112 GGGGTAAATTAAGACAGAGGAGTGAAGTCTTGGGTGTGGTACAGTACACTGCCCCAGCA 5171
|||||
Db 4556 GGGGTAAATTAAGACAGAGGAGTGAAGTCTTGGGTGTGGTACAGTACACTGCCCCAGCA 4615
|||||
OY 5172 GGCCTGTGGCTGTAGAAAAAGTTCAAGGCTTAAGCCCGGACAGCTGCTCAGCTGTGA 5231
|||||
Db 4616 GGCCTGTGGCTGTAGAAAAAGTTCAAGGCTTAAGCCCGGACAGCTGCTCAGCTGTGA 4675
|||||
OY 5232 ATCCACACATTTGGGAGGCGCGAGGCGGGTGTGATCAGAGGTCAAGAGATCTGACCATC 5291
|||||
Db 4676 ATCCACACATTTGGGAGGCGCGAGGCGGGTGTGATCAGAGGTCAAGAGATCTGACCATC 4735
|||||
OY 5292 CTGGCTTAACAGGTTGAACCCCGTCTCTACTTAATAAATTAACAAAAATTTGGCCGGCATGCT 5351
|||||
Db 4736 CTGGCTTAACAGGTTGAACCCCGTCTCTACTTAATAAATTAACAAAAATTTGGCCGGCATGCT 4795
|||||
OY 5352 GGGGGGCACTGTAGTTCCAGCTACTCGGAGGCTGAGGAGGAAGATGCGTGAACCCG 5411
|||||
Db 4796 GGGGGGCACTGTAGTTCCAGCTACTCGGAGGCTGAGGAGGAAGATGCGTGAACCCG 4855
|||||
OY 5412 AGAGGCAAGATTGCAAGTGAAGCCGAGATGCGGCACATGCACTCAGCCTGGGGAGCAAGAG 5471
|||||
Db 4856 AGAGGCAAGATTGCAAGTGAAGCCGAGATGCGGCACATGCACTCAGCCTGGGGAGCAAGAG 4915
|||||
OY 5472 CAAAGCTCCATCTGAAAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5531
|||||
Db 4916 CAAAGCTCCATCTGAAAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4975
|||||
OY 5532 TAATTTGTCTACTTACATGACCTTGGGCAAGGCACTTCCCTTCCCTGGCCCAAGTCAAG 5591
|||||
Db 4976 TAATTTGTCTACTTACATGACCTTGGGCAAGGCACTTCCCTTCCCTGGCCCAAGTCAAG 5035
|||||
OY 5592 GGTGGAATGCACTCCAAAGTCCCTTCCAGCATTAACGCTGCAATGCTTGAAGTGAAGAA 5651
|||||
Db 5036 GGTGGAATGCACTCCAAAGTCCCTTCCAGCATTAACGCTGCAATGCTTGAAGTGAAGAA 5095
|||||
OY 5652 GATGGGAGATTTCCCTCTCTCACCAGCCCGTGTCCACTTCAAGGTGAAGTGAAGCAAG 5711
|||||
Db 5096 GATGGGAGATTTCCCTCTCTCACCAGCCCGTGTCCACTTCAAGGTGAAGTGAAGCAAG 5155
|||||
OY 5712 GAAATGAGTGTCCCAATCCCGAGTTTCCAAAGCCCTTGGGAGCCCTACTGTCAAGGTG 5771
|||||
Db 5156 GAAATGAGTGTCCCAATCCCGAGTTTCCAAAGCCCTTGGGAGCCCTACTGTCAAGGTG 5215
|||||
OY 5772 TGCAGAGAGAGTGAAGGTCAAGTGAAGCAATGCTTCGAAGAGGTCTTCCCTCAATTCGGG 5831
|||||

```

Db      5216  TGCACGAGAGGTGAAGTCAGGTGAGCCAAATCGCCTGGAAGGGTCTGGCCATTCGGG  5275
Oy      5832  ACAGACATCCGGTTCCTCTGCTGTACCAGGATTCCTAGGGCTTTAGCCGATGATCA  5891
Db      5276  ACAGACATCCGGTTCCTCTGCTGTACCAGGATTCCTAGGGCTTTAGCCGATGATCA  5335
Oy      5892  TGGGGGGGGGGGGTTCCTGGGGAGTTCACGCTAATCACTTGGGACAGACGCT  5951
Db      5336  TGGGGGGGGGGGGTTCCTGGGGAGTTCACGCTAATCACTTGGGACAGACGCT  5395
Oy      5952  GGAACCTTCGATGTGCTTATCCAGTGTGGGGGGGACACAGACCAAGACCAATGTC  6011
Db      5396  GGAACCTTCGATGTGCTTATCCAGTGTGGGGGGGACACAGACCAAGACCAATGTC  5455
Oy      6012  CTATCTCAGTGTAGGGGCTCAGAGAGTCTCCACAGACGACGCTCCGAGAGTTGGG  6071
Db      5456  CTATCTCAGTGTAGGGGCTCAGAGAGTCTCCACAGACGACGCTCCGAGAGTTGGG  5515
Oy      6072  GTAGGAATGGAGCAACAGGCTTTTCTCTTGAATTTGGGGGCTGGGGGA  6131
Db      5516  GTAGGAATGGAGCAACCA- GCTTCTTTTCTCTTGAATTTGGGGGCTGGGGGA  5574
Oy      6132  CAGGCTTGAAGATCCCAAGAGAGGGGCAAGAGACT- CCCCCAAGTCTGCAGAG  6190
Db      5575  CAGGCTTGAAGATCCCAAGAGAGGGGCAAGAGACTGCCCGGCAAGTCTGCAGAG  5634
Oy      6191  CGAAGAGGAGAGACCCGACATCAGTGCACCTCCCCACAGGC  6233
Db      5635  C-AGAGAGGAGAGACCCGACATCAGTGCACCTCCCCACAGGC  5676

```

RESULT 6
US-60-466-412-86425

```

: Sequence 86425, Application US/60466412
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: IAKUBOVA, Olga
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001466
: CURRENT APPLICATION NUMBER: US/60/466,412
: CURRENT FILING DATE: 2003-04-30
: NUMBER OF SEQ ID NOS: 429241
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 86425
: LENGTH: 14381
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-466-412-86425

```

Query Match 88.7%; Score 5532.6; DB 101; Length 14381;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 5655; Conservative 0; Mismatches 14; Indels 14; Gaps 10;

```

Oy      558  AAAAAAGCTGTGTTATGTCCTGTCATATCACTACTATGTAATATGTTGCA  617
Db      1  AAAAAAGCTGTGTTATGTCCTGTCATATCACTACTATGTAATATGTTGCA  57
Oy      618  ACTCAAGATCCAGATAGTCAATTTTATGAGCTGTGGCCGTATGCTGTGCACAT  677
Db      58  ACTCAAGATCCAGATAGTCAATTTTATGAGCTGTGGCCGTATGCTGTGCACAT  117
Oy      678  CACTGTGCTGTCTTCTTACACAAAGACGCTATAAACAATATACATCATGAATTTT  737
Db      118  CACTGTGCTGTCTTCTTACACAAAGACGCTATAAACAATATACATCATGAATTTT  177
Oy      738  ATAGCATGAGATTTGAATTTCAATATATTTTATATTAATATATTTTAA  797
Db      178  ATAGCATGAGATTTGAATTTCAATATATTTTATATTAATATATTTTAA  237
Oy      798  AATTTCCCTAACATTTAAAGTAAAGCCGGCA- - -GCCGCCATGCTACAGGC  854

```

```

Db      238  AATTTCCCTAACATTTAAAGTAAAGCCGGCCAGGCGCCCATGCTCAGCC  297
Oy      855  TGTAAATCCAGACATTTGGAGGCTTAGGTGGGACATATCACTTGAATCAACAGTTCAG  914
Db      298  TGTAAATCCAGACATTTGGAGGCTTAGGTGGGACATATCACTTGAATCAACAGTTCAG  357
Oy      915  ACCAGCTGCCCAACATATCAAAACCCATTTCTACTTAAATAATTAATTAATTAATTA  974
Db      358  ACCAGCTGCCCAACATATCAAAACCCATTTCTACTTAAATAATTAATTAATTAATTA  417
Oy      975  ATAGTGTGCACACCTGTGATCCAGACTCTTGGAGGCTGAGGACGAGAAATCGTTGA  1034
Db      418  ATAGTGTGCACACCTGTGATCCAGACTCTTGGAGGCTGAGGACGAGAAATCGTTGA  477
Oy      1035  ACCGTGGAAAGCGAGGTTGACATGACCAACATCAATCCACATCCAGCTCGGGTGA  1094
Db      478  ACCGTGGAAAGCGAGGTTGACATGACCAACATCAATCCAGCTCGGGTGA  537
Oy      1095  CAGAGTGAGACTTCGTCTCAAGCAAAAGTGTAAAGCCATTCCTAATTCAGTGT  1154
Db      538  CAGAGTGAGACTTCGTCTCAAGCAAAAGTGTAAAGCCATTCCTAATTCAGTGT  597
Oy      1155  ACATCAGTGTACTACTCAGGTCTGCTACTCTGCTGAGGCATACCTGAGAAATGA  1214
Db      598  ACATCAGTGTACTACTCAGGTCTGCTACTCTGCTGAGGCATACCTGAGAAATGA  657
Oy      1215  GTTGTGTGTACAGAGACATATTCACATTTTCAATTAATAGACATCAAGTGGCATC  1274
Db      658  GTTGTGTGTACAGAGACATATTCACATTTTCAATTAATAGACATCAAGTGGCATC  717
Oy      1275  CAAGAGG- -TTTTTTTTTACAATCTACACTCCCGCCAGCAACAAATGAGAGTTACTCCA  1333
Db      718  CAAGAGGTTTTTTTTTTACAATCTACACTCCCGCCAGCAACAAATGAGAGTTACTCCA  777
Oy      1334  GATCCTTTAACAAGATGCTTAAAGCCAGTACAGATGAAGAAAGAGTGGAGGGAA  1393
Db      778  GATCCTTTAACAAGATGCTTAAAGCCAGTACAGATGAAGAAAGAGTGGAGGGAA  837
Oy      1394  GCTGCCAGCCCTTCAACATGAAGAAATACCTGTGATAGCCTTGTGATGTGGAAAG  1453
Db      838  GCTGCCAGCCCTTCAACATGAAGAAATACCTGTGATAGCCTTGTGATGTGGAAAG  897
Oy      1454  ATGAATTAAGGGGGTCTGAGACCTGCCCCGTGACATCACTGATCTGAGACCTC  1513
Db      898  ATGAATTAAGGGGGTCTGAGACCTGCCCCGTGACATCACTGATCTGAGACCTC  957
Oy      1514  CAGTCCAGTCTAGCCCCCATGTGTCAATGAGCCAGTGAATGAGCCCTCACTCTGTTTG  1573
Db      958  CAGTCCAG-CTCAGCCCCCATGTGTCAATGAGCCAGTGAATGAGCCCTCACTCTGTTTG  1016
Oy      1574  GTCTTATTTCT- -CCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTATTCAGATGTACA  1632
Db      1017  GTCTTATTTCTCCCATGTGGGGCTGAAGTCTGGATTGAGCCGTTATTCAGATGTACA  1076
Oy      1633  GCTTCTTGACAGGAAGTATGTCTACAAACAGACGAGGGCTTGGCAAGATGATCTAAC  1692
Db      1077  GCTTCTTGACAGGAAGTATGTCTACAAACAGACGAGGGCTTGGCAAGATGATCTAAC  1136
Oy      1693  TGCAAATCTTACTGTGCTGACGACACAGAGTGTCTGTATCTTGAACAAGTTTTCAC  1752
Db      1137  TGCAAATCTTACTGTGCTGACGACACAGAGTGTCTGTATCTTGAACAAGTTTTCAC  1196
Oy      1753  TTCTGTGAGGCAATCCCTTGGCTACAAACACACAGTGTGTGACAGAGTGAATGACG- A  1811
Db      1197  TTCTGTGAGGCAATCCCTTGGCTACAAACACACAGTGTGTGACAGAGTGAATGACGAA  1256
Oy      1812  AGTCCCTTACCTGTAATCCAGACCTTTGGAGAGCCCAAGGGGCTGATGGCTGAGC  1871
Db      1257  AGTCCCTTACCTGTAATCCAGACCTTTGGAGAGCCCAAGGGGCTGATGGCTGAGC  1316
Oy      1872  CTGAGAGGTGACAGATGCGGCGAGTCTCAGAGCCCTGTTGCTGCTGCGGCGCTCTC  1931
Db      1317  CTGAGAGGTGACAGATGCGGCGAGTCTCAGAGCCCTGTTGCTGCTGCGGCGCTCTC  1376

```

QY	1932	TGCGGGGCTCCCACTTGGTGGGCACTTGAGAGGCGCTTCAAGCCCAACCGCTGCACCTGG	1991
Db	1377	TGCGTGGGCTCCCACTTGGTGGGCACTTGAGAGGCGCTTCAAGCCCAACCGCTGCACCTGG	1436
QY	1992	GAGCCCTTTCTTGGGCTGGCCAAAGCCAGAGCCGGCTCCCTCAGCTTGCAGAGAGGTGTG	2051
Db	1437	GAGCCCTTTCTTGGGCTGGCCAAAGCCAGAGCCGGCTCCCTCAGCTTGCAGAGAGGTGTG	1496
QY	2052	GAGGAGAGGCTCAAGCAGAGAACCGGGGCTCGCAGCGGGCTTGGGGCCAGCTGAGATT	2111
Db	1497	GAGGAGAGGCTCAAGCAGAGAACCGGGGCTCGCAGCGGGCTTGGGGCCAGCTGAGATT	1556
QY	2112	CCGGGTGGGCTTGGGCTTGGGGGCCCCGCACTGGGACACGGGGCCAGCCCTGCCAGGC	2171
Db	1557	CCGGGTGGGCTTGGGCTTGGGGGCCCCGCACTGGGACACGGGGCCAGCCCTGCCAGGC	1616
QY	2172	CCCGGGCAATGAGAGGCTTACACACCCGGGCGCAGCGGCTGGGAGGTGTACTGGGTGCC	2231
Db	1617	CCCGGGCAATGAGAGGCTTACACACCCGGGCGCAGCGGCTGGGAGGTGTACTGGGTGCC	1676
QY	2232	CAGCAGTGGCAGCCCGCCGGCGCTGTGCTGTGCTGATTTCTCAGTGGGCTTACAGGCT	2291
Db	1677	CAGCAGTGGCAGCCCGCCGGCGCTGTGCTGTGCTGATTTCTCAGTGGGCTTACAGGCT	1736
QY	2292	TCCCGGGGCGAGGGCTCGGGACCTGAGCGCCCGCATGCGCTGAGCGCTCCCTCCATGGGC	2351
Db	1737	TCCCGGGGCGAGGGCTCGGGACCTGAGCGCCCGCATGCGCTGAGCGCTCCCTCCATGGGC	1796
QY	2352	TCTGTGTGGGCTCCGAGGCTCCCGACGACACACCCCTGTCTCACAGCCCGCAGTCCC	2411
Db	1797	TCTGTGTGGGCTCCGAGGCTCCCGACGACACACCCCTGTCTCACAGCCCGCAGTCCC	1856
QY	2412	ATGCACCAACGAGGCTGAGAGTGGGGGCGCACGGCACCGGAGCTGGGACGAGCTAC	2471
Db	1857	ATGCACCAACGAGGCTGAGAGTGGGGGCGCACGGCACCGGAGCTGGGACGAGCTAC	1916
QY	2472	CCCTGCAGCCCTGGTGGGGAATCCACTGGGTGAAGCAAGCTGGGCTCCTGATGTGGTG	2531
Db	1917	CCCTGCAGCCCTGGTGGGGAATCCACTGGGTGAAGCAAGCTGGGCTCCTGATGTGGTG	1976
QY	2532	AGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCTGAATATACCAATCGACACCTGT	2591
Db	1977	AGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCTGAATATACCAATCGACACCTGT	2036
QY	2592	GTCAGTCAAGGCTCTGTGAATGCACCAATCCACACTCTGTATCTACTACTGTATGGG	2651
Db	2037	GTCAGTCAAGGCTCTGTGAATGCACCAATCCACACTCTGTATCTACTACTGTATGGG	2096
QY	2652	GCCTTGGAGAACCTTTATGTCTAGCTCAGGGATCTGAATATACCAATCGGCACTGTGA	2711
Db	2097	GCCTTGGAGAACCTTTATGTCTAGCTCAGGGATCTGAATATACCAATCGGCACTGTGA	2156
QY	2712	TCTAGTCAAGGTTTGTAAACACACCAATCAGACCCCTGTGTAGTCAAGGATGTGA	2771
Db	2157	TCTAGTCAAGGTTTGTAAACACACCAATCAGACCCCTGTGTAGTCAAGGATGTGA	2216
QY	2772	ATGCACCAATCGACAGCTGTATGTGGCTACTTTCAATGGGCAATCCGTGTGAAGACAC	2831
Db	2217	ATGCACCAATCGACAGCTGTATGTGGCTACTTTCAATGGGCAATCCGTGTGAAGACAC	2276
QY	2832	CAACAGGCTTGTGTAGCAATTAAGCTTCTATCAGCTGGGTGAGGTGGGCTGAGTGC	2891
Db	2277	CAACAGGCTTGTGTAGCAATTAAGCTTCTATCAGCTGGGTGAGGTGGGCTGAGTGC	2336
QY	2892	GAAAGAGATCAGCGAAGGAGATTAAGGTGGGCGCTTTTATAGATTGGGTAGTA	2951
Db	2337	GAAAGAGATCAGCGAAGGAGATTAAGGTGGGCGCTTTTATAGATTGGGTAGTA	2396
QY	2952	AAGGAAATTAAGTCAAAAGGGGTTGTCTCTGGGCGCAGAGTGGGGGTCCGACG	3011
Db	2397	AAGGAAATTAAGTCAAAAGGGGTTGTCTCTGGGCGCAGAGTGGGGGTCCGACG	2456

QY	3012	GTGCTCACTGGGGGTGCTTTTGTAGCCAGAGATGAGCCAGAGAAAGGACTTTCACAGTA	3071
Db	2457	GTGCTCACTGGGGGTGCTTTTGTAGCCAGAGATGAGCCAGAGAAAGGACTTTCACAGTA	2516
QY	3072	ATGTCATTAATTAAGGCAAGGACCCGCCATTTACACCTTTTGTGGTGAATGCATCA	3131
Db	2517	ATGTCATTAATTAAGGCAAGGACCCGCCATTTACACCTTTTGTGGTGAATGCATCA	2576
QY	3132	GTTAAGTTGGGCGACGAGATATTCATCTTTTGTGATCTTTCAGTACTTCAAGCCATC	3191
Db	2577	GTTAAGTTGGGCGACGAGATATTCATCTTTTGTGATCTTTCAGTACTTCAAGCCATC	2636
QY	3192	TGGGGTATATGTCAAGTTTACAGGGAGTGGGATGGCTTGGCTTGGCTCAGAGCTTGA	3251
Db	2637	TGGGGTATATGTCAAGTTTACAGGGAGTGGGATGGCTTGGCTTGGCTCAGAGCTTGA	2696
QY	3252	CAGTACTCTGTGGGGCTTGGAGAAATTTGTGTGACACTGTGTATCTAATGTAATCT	3311
Db	2697	CAGTACTCTGTGGGGCTTGGAGAAATTTGTGTGACACTGTGTATCTAATGTAATCT	2756
QY	3312	AGTGGGACGTGGAGAACCTTTGTGTAGCTCAGGGATTTGTAAACGCCAATCAGCGC	3371
Db	2757	AGTGGGACGTGGAGAACCTTTGTGTAGCTCAGGGATTTGTAAACGCCAATCAGCGC	2816
QY	3372	CCTGTCAAAACAGACACTCGGCTCTACCAATCAGCAGATGTGGGTGGCCAGATTAAG	3431
Db	2817	CCTGTCAAAACAGACACTCGGCTCTACCAATCAGCAGATGTGGGTGGCCAGATTAAG	2876
QY	3432	AGAATTAAGAGAGCTGCCGACGACAGAGTGGGACGACAGGCTCCATACACANT	3491
Db	2877	AGAATTAAGAGAGCTGCCGACGACAGAGTGGGACGACAGGCTCCATACACANT	2936
QY	3492	ATGGCAGCTTTGTCTTTTGTGTGGATTAATCTTGTACTACTGCTTTTGGGTC	3551
Db	2937	ATGGCAGCTTTGTCTTTTGTGTGGATTAATCTTGTACTACTGCTTTTGGGTC	2996
QY	3552	CACACTGTTTTATGAGTGTAAACATCAGCAGCAAGGCTTGCAGCTTCACTCCTGAAGC	3611
Db	2997	CACACTGTTTTATGAGTGTAAACATCAGCAGCAAGGCTTGCAGCTTCACTCCTGAAGC	3056
QY	3612	CACTAAGACACGAGCCACCGGAGAGAAATGAACACTCGGGCGGCTGCTTAAAGGC	3671
Db	3057	CACTAAGACACGAGCCACCGGAGAGAAATGAACACTCGGGCGGCTGCTTAAAGGC	3116
QY	3672	TATTAACACTCACCGCGAAGTCTCAGACTTCACTCTCAGCAGCAGAGCCAGCAACCA	3731
Db	3117	TATTAACACTCACCGCGAAGTCTCAGACTTCACTCTCAGCAGCAGAGCCAGCAACCA	3176
QY	3732	CCAGAAAGAAATCTCGGAACATCTGAACATCAGAAAGAACAAACTCCAGTGCACC	3791
Db	3177	CCAGAAAGAAATCTCGGAACATCTGAACATCAGAAAGAACAAACTCCAGATGCACC	3236
QY	3792	ACCTTAAGAGCTGTAAACACTCAGCGAGGTCCGGGCTTCTTCTTGAATCAGTGAAG	3851
Db	3237	ACCTTAAGAGCTGTAAACACTCAGCGAGGTCCGGGCTTCTTCTTGAATCAGTGAAG	3296
QY	3852	ACCAAGCACTCACAGTTTGGGACACAAAGCCAGAGTTTGAATCAGGCTGGGCAACAT	3911
Db	3297	ACCAAGCACTCACAGTTTGGGACACAAAGCCAGAGTTTGAATCAGGCTGGGCAACAT	3356
QY	3912	GATGAATGGCCCTCTCGCAAAAAAATTAACAAAAATTTGGGGGAGCATGTGGT	3971
Db	3357	GATGAATGGCCCTCTCTCGCAAAAAAATTAACAAAAATTTGGGGGAGCATGTGGT	3416
QY	3972	CCGTGCTGTGTGCCAGCTACCGGGGAGGCTTAAGTGGAGATCGCTTGAAGCTGGGA	4031
Db	3417	CCGTGCTGTGTGCCAGCTACCGGGGAGGCTTAAGTGGAGATCGCTTGAAGCTGGGA	3476
QY	4032	GGTGAAGCTTGCAGTGAAGTGTGATTTACACAGCCCTCTAGGCTGGGGAGACAGCTGA	4091
Db	3477	GGTGAAGCTTGCAGTGAAGTGTGATTTACACAGCCCTCTAGGCTGGGGAGACAGCTGA	3536
QY	4092	GACCTGTTCCTCCCTCCGCAAAAAAATTTGACAAAAAGTGAATTAAGAGTCCCTGATGTGG	4151

|||||
Db 3537 GACCTGTTCCCTCCGCAAAAAATTGACAAAAGTATTAAGAGGTGCTGATATAGG 3596
Qy 4152 CTAGGCCAGTGGCTCATGCTGTAAATCCAGCACTTTGGAGACCCGAGCGGGGCTC 4211
Db 3597 CTAGGCCAGTGGCTCATGCTGTAAATCCAGCACTTTGGAGACCCGAGCGGGGCTC 3656
Qy 4212 ACCTAAGGTGAGGTGAGAACGAGCTGGCCAAACATGGAGAAAGCCATCTCTTCAA 4271
Db 3657 ACCTAAGGTGAGGTGAGAACGAGCTGGCCAAACATGGAGAAAGCCATCTCTTCAA 3716
Qy 4272 AAATACAAATTAACCCGCTGTGGGGCAGTGTGAGCATGCTGTAAATCCAGCTACT 4331
Db 3717 AAATACAAATTAACCCGCTGTGGGGCAGTGTGAGCATGCTGTAAATCCAGCTACT 3776
Qy 4332 CAGAGGCTGAGGAGAGAGAAATCACTTGAACCCAGAGGCGGGCTTGCAGTAGCCGAG 4391
Db 3777 CAGAGGCTGAGGAGAGAGAAATCACTTGAACCCAGAGGCGGGCTTGCAGTAGCCGAG 3836
Qy 4392 ATCGTGCATTTGCACTCCACCACTCCAGCTGGGCAACAGAGCCAAACTCTGTAA 4451
Db 3837 ATCGTGCATTTGCACTCCACCACTCCAGCTGGGCAACAGAGCCAAACTCTGTAA -A 3895
Qy 4452 AAAAAAAAAAAAAAGTCCCTGACATATAGAGGTGTGCAATGCAATAGTTGCCAGCAA 4511
Db 3896 AAAAAAAAAAAAAAGTCCCTGACATATAGAGGTGTGCAATGCAATAGTTGCCAGCAA 3955
Qy 4512 CATGTTAAAGAAATGAGAGCTCCCTGCTTCATGCTGTAAATCCCACTTCTTAAAG 4571
Db 3956 CATGTTAAAGAAATGAGAGCTCCCTGCTTCATGCTGTAAATCCCACTTCTTAAAG 4015
Qy 4572 CAGAGTGCAGTGTGCTCATGCTTAAATCCAGCACTTTGGAGGCGGAGGGGTGATC 4631
Db 4016 CAGAGTGCAGTGTGCTCATGCTTAAATCCAGCACTTTGGAGGCGGAGGGGTGATC 4075
Qy 4632 ACCTGAGTGCAGAGTTCGAGACCAAGCTGACCAACAATGTTAAATCCCACTTAC 4691
Db 4076 ACCTGAGTGCAGAGTTCGAGACCAAGCTGACCAACAATGTTAAATCCCACTTAC 4135
Qy 4692 TAAAAATATCAAAATTAATAGATGATGATGGTGTGATGCTGATTAATCCCACTTCTTAAAG 4751
Db 4136 TAAAAATATCAAAATTAATAGATGATGATGGTGTGATGCTGATTAATCCCACTTCTTAAAG 4195
Qy 4752 GCTGAGGAGGAAATTAATAGAAACAGAGGAGGAGGTTGTAGTGAGCCGAGATCTG 4811
Db 4196 GCTGAGGAGGAAATTAATAGAAACAGAGGAGGAGGTTGTAGTGAGCCGAGATCTG 4255
Qy 4812 CCATTCAGTTCAGAGCTGAGCAATGAGCGAAATCTCATCTCAAAAAACAACAAAAAA 4871
Db 4256 CCATTCAGTTCAGAGCTGAGCAATGAGCGAAATCTCATCTCAAAAAACAACAAAAAA 4315
Qy 4872 CCCAGTCTCTATCCCAAGGAGGAGTGGGTACAGAGCTGGGCCACATCACTGCAAGTGTG 4931
Db 4316 CCCAGTCTCTATCCCAAGGAGGAGTGGGTACAGAGCTGGGCCACATCACTGCAAGTGTG 4375
Qy 4932 AGCCACAGAGCTAAGGCGAGAGCTCAGAGACCGGAGACAGATAACAGTGTGATCAG 4991
Db 4376 AGCCACAGAGCTAAGGCGAGAGCTCAGAGACCGGAGACAGATAACAGTGTGATCAG 4435
Qy 4992 TGTGTGAGATCAGACGTCCCTGCTGCAATGTGTACCAACAGAGGCGGCCCAAGCAGAG 5051
Db 4436 TGTGTGAGATCAGACGTCCCTGCTGCAATGTGTACCAACAGAGGCGGCCCAAGCAGAG 4455
Qy 5052 TGGCCCATCCAGTACACACATCCATCTTCATCCAGAGATGTCTGTTTCTTGGCAGCT 5111
Db 4496 TGGCCCATCCAGTACACACATCCATCTTCATCCAGAGATGTCTGTTTCTTGGCAGCT 4555
Qy 5112 GGGGTAAATTAAGGACAGAGGTGACATCTTGGTGTGATCAGTCAAGTGTGAGGCA 5171
Db 4556 GGGGTAAATTAAGGACAGAGGTGACATCTTGGTGTGATCAGTCAAGTGTGAGGCA 4615
Qy 5172 GGCTTGTGCTGTAGAAAAAGTTCAGGCTTAGGCGGAGCAGTGGCTCAGCGCTGA 5231
|||||

Db 4616 GGCTTGTGCTGTAGAAAAAGTTCAGGCTTAGGCGGAGCAGGTTGGCTCAGCGCTGA 4675
Qy 5232 ATCCAGACATTTGGGAGCGGAGCGGGTGGATACAGAGTCAAGAGATCTGACATC 5291
Db 4676 ATCCAGACATTTGGGAGCGGAGCGGGTGGATACAGAGTCAAGAGATCTGACATC 4735
Qy 5292 CTGGCTAACAGGTGAAACCCCGTCTCTACTTAAATACAAAAATTTGGCGGGGATGT 5351
Db 4736 CTGGCTAACAGGTGAAACCCCGTCTCTACTTAAATACAAAAATTTGGCGGGGATGT 4795
Qy 5352 GCGGCGACCTGTACTTCCAGTACTCGGAGGCTGAGGACAGAGATGCGCTGAACCG 5411
Db 4796 GCGGCGACCTGTACTTCCAGTACTCGGAGGCTGAGGACAGAGATGCGCTGAACCG 4855
Qy 5412 AGAGCAAGTGTGAGTGAAGCCGAGATCGCGCCACTGCACTCCAGCTGGCGCAGAG 5471
Db 4856 AGAGCAAGTGTGAGTGAAGCCGAGATCGCGCCACTGCACTCCAGCTGGCGCAGAG 4915
Qy 5472 CAAGACTCATCTGAAAAAGAAAAAGAAACCTTCAGGCTTAGGCGCAGAGCGCGTG 5531
Db 4916 CAAGACTCATCTGAAAAAGAAAAAGAAACCTTCAGGCTTAGGCGCAGAGCGCGTG 4975
Qy 5532 TAATTCGTCACTTACATGACCTTGGCAAGGCACTTCCTGCGCCAGTTACGG 5591
Db 4976 TAATTCGTCACTTACATGACCTTGGCAAGGCACTTCCTGCGCCAGTTACGG 5035
Qy 5592 GGTGGATTCGACTCCAGTCCCTTCCAGCATTAACCTGCATGTTCTAAGATGAGA 5651
Db 5036 GGTGGATTCGACTCCAGTCCCTTCCAGCATTAACCTGCATGTTCTAAGATGAGA 5095
Qy 5652 GATGGGAGAGTTTCCCTCTCTCACCCAGCCCGGTGCTCAAGGTGAATGACCAAG 5711
Db 5096 GATGGGAGAGTTTCCCTCTCTCACCCAGCCCGGTGCTCAAGGTGAATGACCAAG 5155
Qy 5712 GAAGTCACTGTCCTCAATCCCGACTTCCAAAGCCCTTGGGAGCCCTACTGTCAAGGTCG 5771
Db 5156 GAAGTCACTGTCCTCAATCCCGACTTCCAAAGCCCTTGGGAGCCCTACTGTCAAGGTCG 5215
Qy 5772 TGACAGAGAGGTGAAGTCAAGTGAAGTCCGCTCGAAGAGGTCCTGCTCATTCGG 5831
Db 5216 TGACAGAGAGGTGAAGTCAAGTGAAGTCCGCTCGAAGAGGTCCTGCTCATTCGG 5275
Qy 5832 ACAGACATCCGTTTCTCTGCTTACCGGATCTAGGGCTTTAGCCGAATGATCA 5891
Db 5276 ACAGACATCCGTTTCTCTGCTTACCGGATCTAGGGCTTTAGCCGAATGATCA 5335
Qy 5892 TGGGGGCGGGGGGTTTCTGGGGAGTTCCAGCTAATCAACTTGGGACAGAGACCT 5951
Db 5336 TGGGGGCGGGGGGTTTCTGGGGAGTTCCAGCTAATCAACTTGGGACAGAGACCT 5395
Qy 5952 GGAATCTTCGATGGTGCCTATCCAGTGTGGGGTGGGACAGACCAAGCAATGTC 6011
Db 5396 GGAATCTTCGATGGTGCCTATCCAGTGTGGGGTGGGACAGACCAAGCAATGTC 5455
Qy 6012 CTTATCTCAGGTAGGGGCTCAGAGAGTTCACAGACAGGACCTCCGAGAGTTTGGG 6071
Db 5456 CTTATCTCAGGTAGGGGCTCAGAGAGTTCACAGACAGGACCTCCGAGAGTTTGGG 5515
Qy 6072 GTAGGAATGGGAGCAACAGGCTTTTCTTCTCTTGAATTTGGGGGCTTGGGGA 6131
Db 5516 GTAGGAATGGGAGCAACA -GCTTCTTTTCTCTCTTGAATTTGGGGGCTTGGGGA 5574
Qy 6132 CAGGCTTGAATTCACAAAGAGAGGGCAAAAGCACT -CCGCCACAAGTGTGCAAG 6190
Db 5575 CAGGCTTGAATTCACAAAGAGAGGGCAAAAGCACTGCCGCCGCAAGTGTGCAAG 5634
Qy 6191 CGAGAGAGGAGACCCGACATCAGTGTCCACTTCCCAAGAGC 6233
Db 5635 C-AGAGAGGAGACCCGACATCAGTGTCCACTTCCCAAGAGC 5676
|||||

RESULT 7
US-09-845-020a-6

; Sequence 6: Application US/09845020A
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michel W.
; APPLICANT: Seiden, Richard F.
; TITLE OF INVENTION: Genomic Sequences for Protein Production
; TITLE OF INVENTION: and Delivery
; FILE REFERENCE: 50010/017003
; CURRENT APPLICATION NUMBER: US/09/845, 020A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 09/305,384
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: US 60/084,649
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 2834
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-020A-6

Query Match 45.5%; Score 2834; DB 35; Length 2834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1890 CCGGCAATCTCAGACCCCTGTCCTGCGGCGCTCTGCTGGGCTCCACTTC 1949
DB 1 CCGGCAATCTCAGACCCCTGTCCTGCGGCGCTCTGCTGGGCTCCACTTC 60
QY 1950 GGTGGCACTTGAAGAGCCCTTCAAGCCAGCCGCTGAGAGAGCCCTTTTGGGCTG 2009
DB 61 GGTGGCACTTGAAGAGCCCTTCAAGCCAGCCGCTGAGAGAGCCCTTTTGGGCTG 120
QY 2010 GCCAAGCCAGAGAGCCGCTCTCCCTCAGCTTTCAGAGAGTGTGAGAGAGAGCTCAAGCA 2069
DB 121 GCCAAGCCAGAGAGCCGCTCTCCCTCAGCTTTCAGAGAGTGTGAGAGAGAGCTCAAGCA 180
QY 2070 GGAACCGGGGCTGCGACAGCGGCTTCCGGGCAAGCTGAGTCCGGGTGGGCTG 2129
DB 181 GGAACCGGGGCTGCGACAGCGGCTTCCGGGCAAGCTGAGTCCGGGTGGGCTG 240
QY 2130 GGGGGGCGCGCACTGGAGAGAGCGGCGCAAGCCCTTCCAGGCGCGGCAATAGAGCT 2189
DB 241 GGGGGGCGCGCACTGGAGAGAGCGGCGCAAGCCCTTCCAGGCGCGGCAATAGAGCT 300
QY 2190 TAGCACCGGGGCGAGGCTGCGAGAGTGTACTGGGTGCCCGCAGCAGTGCAGCCGCC 2249
DB 301 TAGCACCGGGGCGAGGCTGCGAGAGTGTACTGGGTGCCCGCAGCAGTGCAGCCGCC 360
QY 2250 GGGGCTGTGCTGCTGATTTCTCACTGGGCTTACAGGCTTCCCGGGGCGAGGCTC 2309
DB 361 GGGGCTGTGCTGCTGATTTCTCACTGGGCTTACAGGCTTCCCGGGGCGAGGCTC 420
QY 2310 GGGACCTGAGAGCCCGCATGCTGAGCTCCCTCCATGGGCTGCTGTGGGCGCAGGC 2369
DB 421 GGGACCTGAGAGCCCGCATGCTGAGCTCCCTCCATGGGCTGCTGTGGGCGCAGGC 480
QY 2370 TCCCGAGAGAGCAACCCCTGCTCCAGAGCGCCCAAGTCCCATGCAACAGCAAGGCT 2429
DB 481 TCCCGAGAGAGCAACCCCTGCTCCAGAGCGCCCAAGTCCCATGCAACAGCAAGGCT 540
QY 2430 GAGAACTGGGGGCGAGCGGAGCTGAGAGAGTACAGGCTTCCCGGCGAGGCTGCTG 2489
DB 541 GAGAACTGGGGGCGAGCGGAGCTGAGAGAGTACAGGCTTCCCGGCGAGGCTGCTG 600
QY 2490 GATCCACTGGGAGAGGAGCTGGGCTCTGAGTGTGGTGAAGTGTGAGAACTTTA 2549
DB 601 GATCCACTGGGAGAGGAGCTGGGCTCTGAGTGTGGTGAAGTGTGAGAACTTTA 660
QY 2550 TGTCTAGCTCAGAGGATGTAATACACCAATCAGACCTGTGTCTAGCTCAGGCTCTGT 2609
DB 661 TGTCTAGCTCAGAGGATGTAATACACCAATCAGACCTGTGTCTAGCTCAGGCTCTGT 720

QY 2610 GAATGCACCAATCCACTCTGTATCTAGCTACTCTGATGGGCTTGGAAACCTTTAT 2669
DB 721 GAATGCACCAATCCACTCTGTATCTAGCTACTCTGATGGGCTTGGAAACCTTTAT 780
QY 2670 GTCTAGCTCAGAGGATGTAATACCAATCCAGGAGCTGTATCTAGCTCAGGCTTGTGA 2729
DB 781 GTCTAGCTCAGAGGATGTAATACCAATCCAGGAGCTGTATCTAGCTCAGGCTTGTGA 840
QY 2730 AACACCAATCAGACACCTGTGTCTAGCTCAGGATGTGTAATGCACCAATGCAGAGTC 2789
DB 841 AACACCAATCAGACACCTGTGTCTAGCTCAGGATGTGTAATGCACCAATGCAGAGTC 900
QY 2790 TGTATCTGCTACTTTCATGAGCATCCGTGTGAAGAGACCAACAGGCTTGTGTGA 2849
DB 901 TGTATCTGCTACTTTCATGAGCATCCGTGTGAAGAGACCAACAGGCTTGTGTGA 960
QY 2850 GCAATTAAGCTTTCATCACCCTGGGTGAGGCTGGGCTGAGTCCCAAAAAGAGTCAGCGAA 2909
DB 961 GCAATTAAGCTTTCATCACCCTGGGTGAGGCTGGGCTGAGTCCCAAAAAGAGTCAGCGAA 1020
QY 2910 GGGAGATTAAGGTGGGCGCTTTTATAGATTTGGTAGGTAAAGAAATTAACAGTCAA 2969
DB 1021 GGGAGATTAAGGTGGGCGCTTTTATAGATTTGGTAGGTAAAGAAATTAACAGTCAA 1080
QY 2970 AGGGGTTTGTCTCTGCGGCGAGAGTGGGGGCTGCAAGGCTCTCACTGGGGTGTCT 3029
DB 1081 AGGGGTTTGTCTCTGCGGCGAGAGTGGGGGCTGCAAGGCTCTCACTGGGGTGTCT 1140
QY 3030 TTTTGAAGCCAGAGTGAAGCCAGAAAGAGATTTTCAAGAGTATGTCATTAATTAAGGCA 3089
DB 1141 TTTTGAAGCCAGAGTGAAGCCAGAAAGAGATTTTCAAGAGTATGTCATTAATTAAGGCA 1200
QY 3090 AGGACCCGCAATTTACACCTCTTTTGTGAGATGTCATCAGTAAAGTTGGGCGAGGCG 3149
DB 1201 AGGACCCGCAATTTACACCTCTTTTGTGAGATGTCATCAGTAAAGTTGGGCGAGGCG 1260
QY 3150 ATATTCATCTTTTGTGATTTCTCACTTTCAGGCTTGGGCTGATGTGCAAG 3209
DB 1261 ATATTCATCTTTTGTGATTTCTCACTTTCAGGCTTGGGCTGATGTGCAAG 1320
QY 3210 TTACAGGGAGTGAAGTGGCTTGGCTGGGCTGAGAGCTTACAGCTACTGCTGGGCG 3269
DB 1321 TTACAGGGAGTGAAGTGGCTTGGCTGGGCTGAGAGCTTACAGCTACTGCTGGGCG 1380
QY 3270 CTTGAGAAATGTTTGTGTCAGACCTGTATCTAGTTAATCTAGTGGGAGCTGAGAAAC 3329
DB 1381 CTTGAGAAATGTTTGTGTCAGACCTGTATCTAGTTAATCTAGTGGGAGCTGAGAAAC 1440
QY 3330 CTTTGTGTAGCTCAGAGGATTTTAAACGCAATCAGGCGCTTCAAAACAGACAC 3389
DB 1441 CTTTGTGTAGCTCAGAGGATTTTAAACGCAATCAGGCGCTTCAAAACAGACAC 1500
QY 3390 TCGGCTCTCAATCAGCAGAGATGGGAGGCGCAGATTAAGAAATTAAGAGAGGCTGC 3449
DB 1501 TCGGCTCTCAATCAGCAGAGATGGGAGGCGCAGATTAAGAAATTAAGAGAGGCTGC 1560
QY 3450 CCGAGCAGAGAGTGGCAACGCGCAGAGTCCATTCACCAATATGCGAGCTTGTCTTT 3509
DB 1561 CCGAGCAGAGAGTGGCAACGCGCAGAGTCCATTCACCAATATGCGAGCTTGTCTTT 1620
QY 3510 TGTCTTTTGGGATTAATCTTGTCTAGCTGCTGCTTTTGGGCTCAACACTGCTTTATGAGC 3569
DB 1621 TGTCTTTTGGGATTAATCTTGTCTAGCTGCTGCTTTTGGGCTCAACACTGCTTTATGAGC 1680
QY 3570 TGTAACTACACAGAGAGTGTGAGCTTCACTCTGAAAGCCACTTAAGCCAGGAGCC 3629
DB 1681 TGTAACTACACAGAGAGTGTGAGCTTCACTCTGAAAGCCACTTAAGCCAGGAGCC 1740
QY 3630 ACCGGAGGAATGAACAACTCCGCGCGGCTGCTTAAAGAGCTTAACACTCACCGGAA 3689
DB 1741 ACCGGAGGAATGAACAACTCCGCGCGGCTGCTTAAAGAGCTTAACACTCACCGGAA 1800

3690 GGTCTGACGCTTCACTCCCTCAGCCAGAGACACAGAACCCACCGAAGAGAAACTGC 3749
1801 GGTCTGACGCTTCACTCCCTCAGCCAGAGACACAGAACCCACCGAAGAGAAACTGC 1860
3750 GAACACATCTGACATCAGAAAGAAACAACTCCAGATCCACACCTTAAAGCTGTACA 3809
1861 GAACACATCTGACATCAGAAAGAAACAACTCCAGATCCACACCTTAAAGCTGTACA 1920
3810 CTCACCTGGAGGCTCCGCGCTCTTCTTGAAGTCAGTGAAGACCAACACTCAGCAATT 3869
1921 CTCACCTGGAGGCTCCGCGCTCTTCTTGAAGTCAGTGAAGACCAACACTCAGCAATT 1980
3870 TCGGACACAGCCCAAGAGATTGAGATCAGCTGGGCAACATGATGAATGCCCTCTGTG 3929
1981 TCGGACACAGCCCAAGAGATTGAGATCAGCTGGGCAACATGATGAATGCCCTCTGTG 2040
3930 CAAAAAATTTACAAAAATTTGCGGAGCATGCTGCTGCCCTGTGCTCCAG 3989
2041 CAAAAAATTTACAAAAATTTGCGGAGCATGCTGCTGCCCTGTGCTCCAG 2100
3990 CTAGCGGGAGGCTTAAAGTGGAGATGCTTGAAGCTGGAGGCTGAAGACTGCACTGAG 4049
2101 CTAGCGGGAGGCTTAAAGTGGAGATGCTTGAAGCTGGAGGCTGAAGACTGCACTGAG 2160
4050 CTGTGATTTGACACAGCCCTCTAGGCTGGGGGACAGACTGAGACCTGTTCCCTCCG 4109
2161 CTGTGATTTGACACAGCCCTCTAGGCTGGGGGACAGACTGAGACCTGTTCCCTCCG 2220
4110 CAAAAAATTTACAAAAATTTGATTAAGAGTGCCTGATGATGAGGCTGAGGCTCAT 4169
2221 CAAAAAATTTACAAAAATTTGATTAAGAGTGCCTGATGATGAGGCTGAGGCTCAT 2280
4170 GCTGTATTTCCAGACTTTGGGAAAGCGGGGGGGGCTACCTAAGTCAAGAGTGT 4229
2281 GCTGTATTTCCAGACTTTGGGAAAGCGGGGGGGGCTACCTAAGTCAAGAGTGT 2340
4230 GAGACACAGCTTGGCCAAACATGAGAAAGCCCATCTCTTAAAAATTAACAAATTTAGCCGG 4289
2341 GAGACACAGCTTGGCCAAACATGAGAAAGCCCATCTCTTAAAAATTAACAAATTTAGCCGG 2400
4290 CTGTGGGGGAGTGTGAGCATGCTTAAATCCAGTACTCAGAGGCTGAGGACAGA 4349
2401 CTGTGGGGGAGTGTGAGCATGCTTAAATCCAGTACTCAGAGGCTGAGGACAGA 2460
4350 GAATCACTTGAACCCAGAGGCGGGGCTTGAAGTGAAGCGAGATGCTCCATGCACTCC 4409
2461 GAATCACTTGAACCCAGAGGCGGGGCTTGAAGTGAAGCGAGATGCTCCATGCACTCC 2520
4410 ACCCACTCCAGCTTGGCAACAGCAAGCCAACTGTCTTAAAAAATTAACAAATTTAGT 4469
2521 ACCCACTCCAGCTTGGCAACAGCAAGCCAACTGTCTTAAAAAATTAACAAATTTAGT 2580
4470 CCTGACATTAAGAGTGTGCAATGCAATGCTTGGCAGGCAACATGTTTAAAGATGTGA 4529
2581 CCTGACATTAAGAGTGTGCAATGCAATGCTTGGCAGGCAACATGTTTAAAGATGTGA 2640
4530 GCTCCTGCTTCATGCTGCTTAAAAAACCACCTCAAGGCCAGGCTGCTCAT 4589
2641 GCTCCTGCTTCATGCTGCTTAAAAAACCACCTCAAGGCCAGGCTGCTCAT 2700
4590 GCTTATATCCAGCACTTTGGAGGCGGAGGCTGGATCACCCTGAGGCTCAGAGTTC 4649
2701 GCTTATATCCAGCACTTTGGAGGCGGAGGCTGGATCACCCTGAGGCTCAGAGTTC 2760
4650 GAGACACAGCTTGGCAACAGCAACATGCTGAATCCACCTCTCTAAAAATTAACAAATTTAGA 4709
2761 GAGACACAGCTTGGCAACAGCAACATGCTGAATCCACCTCTCTAAAAATTAACAAATTTAGA 2820
4710 TGAGCATGCTGCTG 4723
2821 TGAGCATGCTGCTG 2834

RESULT 8
US-09-760-457-490
; Sequence 490, Application US/09760457
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1215
; CURRENT APPLICATION NUMBER: US/09/760,457
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 490
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1465)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1478)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1548)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1602)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1603)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1818)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1819)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1824)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1828)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1838)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1841)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1875)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2145)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2298)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-760-457-490

Query Match	35.7%	Score 2225.8	DB 33	Length 2409
Best Local Similarity	98.7%	Pred. No. 1.6e-257		
Matches 2367	Conservative 0	Mismatches 14	Indels 17	Gaps 12
QY	3371	CCCTCTCAAAACAGACACTCGGCTCTACCAATCAGCAGATG-TGGGTGGGGCCAGATA	3429	
Db	13	CCNMTCAAAAACAGACACTCGGCTCTACCAATCAGCAGATGNTGGGTGGGGCCAGATA	72	
QY	3430	AGAGATTAAGAGCGGTCGCCGAGCCAGCAGTGGCAACGGCACAGTCCCTATCCACA	3489	
Db	73	AGAGATTAAGAGCGGTCGCCGAGCCAGCAGTGGCAACGGCACAGTCCCTATCCACA	132	
QY	3490	ATATGGCAGCTTTGTTCTTTTGTGTGTGCGATAAATCTTGCTACTGCTGCTTTT	3548	
Db	133	ATATGGCAGCTTTGTTCTTTTGTGTGTGCGATAAATCTTGCTACTGCTGCTTTTNTGG	192	
QY	3549	GTCACACTGCTTTTATGAGCTGTAACTACACACGAAAGGTCTGACCTTCACTCTGA	3608	
Db	193	GTCACACTGCTTTTATGAGCTGTAACTACACACGAAAGGTCTGACCTTCACTCTGA	252	
QY	3609	AGCCACTAAGACACGAGCCGAGGAGGAAATGAACAATCCGGCCGCTGCTTAAAG	3668	
Db	253	AGCCACTAAGACACGAGCCGAGGAGGAAATGAACAATCCGGCCGCTGCTTAAAG	312	
QY	3669	AGCTATAACTCAACCGGAGAGTCTGCAAGCTTCACTCTCAGCCAGCAGACACACGAA	3728	
Db	313	AGCTATAACTCAACCGGAGAGTCTGCAAGCTTCACTCTCAGCCAGCAGACACGAA	372	
QY	3729	CCACCAAGAGAGAACTCGGAAACATCTGAAACATCAGAAAGAACTCCAGATGC	3788	
Db	373	CCACCAAGAGAGAACTCGGAAACATCTGAAACATCAGAAAGAACTCCAGATGC	432	
QY	3789	ACCACTTAAGAGCTGTAACTCACTGCGGAGGTCCGGGCTTCTTCTTGAAGTCACT	3848	
Db	433	ACCACTTAAGAGCTGTAACTCACTGCGGAGGTCCGGGCTTCTTCTTGAAGTCACT	492	
QY	3849	GAGACCAAGCACTCAGAGTTCGGACACAGCCAGAGATTGAGATCAGCTGGGCAA	3908	
Db	493	GAGACCAAGCACTCAGAGTTCGGACACAGCCAGAGATTGAGATCAGCTGGGCAA	552	
QY	3909	CATATGAAATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGGACATGCT	3968	
Db	553	CATATGAAATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGGACATGCT	612	
QY	3969	GGTCGCTGCTGTGCTCCGAGCTACGCGGAGGCTTAAAGTGGAGATCGTTAGCTG	4028	
Db	613	GGTCGCTGCTGTGCTCCGAGCTACGCGGAGGCTTAAAGTGGAGATCGTTAGCTG	672	
QY	4029	GGAGTGAAGACTCAGTGAAGCTGTATTGACACAGCCCTTAGGCTGGGGACAGAC	4088	
Db	673	GGAGTGAAGACTCAGTGAAGCTGTATTGACACAGCCCTTAGGCTGGGGACAGAC	732	
QY	4089	TGAGACCTGTTTCCCTCCGCAAAAAATTTGACAAAAATGTAAATAGAGTGGCTGATA	4148	
Db	733	TGAGACCTGTTTCCCTCCGCAAAAAATTTGACAAAAATGTAAATAGAGTGGCTGATA	792	
QY	4149	TGGCTAGGCCAGTGGCTCTATGCTGTAAATCCAGCACTTTGGAGCCAGCGGGCGG	4208	
Db	793	TGGCTAGGCCAGTGGCTCTATGCTGTAAATCCAGCACTTTGGAGAACCCAGCGGGCGG	852	
QY	4209	GTCACCTAAGAGTGAAGTGTGAGACCAAGCCTGGCCAAACATGGAGAAAGCCATCTTTC	4268	
Db	853	GTCACCTAAGAGTGAAGTGTGAGACCAAGCCTGGCCAAACATGGAGAAAGCCATCTTTC	912	
QY	4269	TAAAAATACAAAAATTAAGCCGGCTGTGGGGCAGTGGTGAAGCATGCTGTAAATCCAGCT	4328	
Db	913	TAAAAATACAAAAATTAAGCCGGCTGTGGGGCAGTGGTGAAGCATGCTGTAAATCCAGCT	972	
QY	4329	ACTCAGAGGCTGAAGGAGAGATCACTTGAACCCAGAGGCGGGGTTGCACTGAGCC	4388	
Db	973	ACTCAGAGGCTGAAGGAGAGATCACTTGAACCCAGAGGCGGGGTTGCACTGAGCC	1032	
QY	4389	GAGATGTCGTCATTTGACATCCACCACCTCAGGCTGGGCAACAAGGCCAAACTCTCTCT	4448	

Db	1033	GAGATCGTCATTCACCTCCACCTACCTCCAGCCTGGGCAACAAGACCAACTCTCTCT	1092	
QY	4449	TAAAAAAGAGAGAGTCCCTGACATTAATGAAGAGTGTGCAATGCAATGATTGGCCAGG	4508	
Db	1093	T-TAAAAAAGAGAGAGTCCCTGACATTAATGAAGAGTGTGCAATGCAATGATTGGCCAGG	1151	
QY	4509	CAACATGTTTAAAGATGTGAGCTCTGCTTCCATGGTCTGTTAAAAACCAACCTCA	4568	
Db	1152	CAACATGTTTAAAGATGTGAGCTCTGCTTCCATGGTCTGTTAAAAACCAACCTCA	1211	
QY	4569	AGGCCAGGTGCAAGGCTCATGCTTAAATCCAGCACTTTGGAGAGCCGAGGGGGTGG	4628	
Db	1212	AGGCCAGGTGCAAGGCTCATGCTTAAATCCAGCACTTTGGAGAGCCGAGGGGGTGG	1271	
QY	4629	ATCACCTGAGTCAAGAGTTCAGACACAGCCTGACCAACAATGTTGTAATCCACCTC	4688	
Db	1272	ATCACCTGAGTTCAGAGATTCAGACACAGCCTGACCAACAATGTTGTAATCCACCTC	1331	
QY	4689	TACTTAAAAATCAAAAATTTAGATGAGCATGTTGTCATGCTGTAAATCCACCTTGG	4748	
Db	1332	TACTTAAAAATCAAAAATTTAGATGAGCATGTTGTCATGCTGTAAATCCACCTTGG	1391	
QY	4749	GAGGCTGAGGAGGAAAAATCACTAGAACACAGGAGCGGAGAGTGTAGTGGCCGAGATC	4808	
Db	1392	GAGGCTGAGGAGGAAAAATCACTAGAACACAGGAGCGGAGAGTGTAGTGGCCGAGATC	1451	
QY	4809	GTCGATTGCACT-CCAGCCTGAGCA-ATGAGCCAAATCCATCTCAAAAAACAACAAC	4866	
Db	1452	GTCGATTGCACTTCCAGCCTGAGCAANATGAGCAAAATCCATCTCAAAAAACAACAAC	1511	
QY	4867	AAAAACCACTCTTACTCCAGGAGAGTGGGTACA-GAGTGGGCCACATCACTGCAAG	4925	
Db	1512	AAAAACCACTCTTACTCCAGGAGAGTGGGTACA-GAGTGGGCCACATCACTGCAAG	1571	
QY	4926	GTCGAGCCACAGAGTAAAGGAGGAGCTG--CAGGACCCGGGACAGATTAACGTGTGT	4983	
Db	1572	GTCGAGCCACAGAGTAAAGGAGGAGCTG--CAGGACCCGGGACAGATTAACGTGTGT	1631	
QY	4984	GAGATCAGTGTGTGAGATCAGACCTCCCTGCTTGGTGAACACAGGGGGCCGCCAAGC	5043	
Db	1632	GAGATCAGTGTGTGAGATCAGACCTCCCTGCTTGGTGAACACAGGGGGCCGCCAAGC	1691	
QY	5044	ACCAAGATGGCCCCATCCAGTCAACCAATCCACTTTCATCCAGAGATGTCTGTTCTT	5103	
Db	1692	ACCAAGATGGCCCCATCCAGTCAACCAATCCACTTTCATCCAGAGATGTCTGTTCTT	1751	
QY	5104	GGCAGGCTGGGTTAATTAGACAGAGGTGACAGTCTTGGGTGTGTCAGTCACTGC	5163	
Db	1752	GGCAGGCTGGGTTAATTAGACAGAGGTGACAGTCTTGGGTGTGTCAGTCACTGC	1811	
QY	5164	CCGAGGCA---GGCTTGTGGCCTGT--AGAAAACTTCAAGGCTTGGGCGGGCACAGGT	5217	
Db	1812	CCGAGGCA---GGCTTGTGGCCTGT--AGAAAACTTCAAGGCTTGGGCGGGCACAGGT	1871	
QY	5218	GGC-TCAGGCTGTAAATCCAGCACTTTGGAGAGCCGAGCGGGGTGATCAGAGGTCAAG	5276	
Db	1872	GGC-TCAGGCTGTAAATCCAGCACTTTGGAGAGCCGAGCGGGGTGATCAGAGGTCAAG	1931	
QY	5277	GAGATGTCACATTCCTGGCTTAACAGGCTTAAACCCGCTTCTTACTTAAAAATCAAAAAA	5336	
Db	1932	GAGATGTCACATTCCTGGCTTAACAGGCTTAAACCCGCTTCTTACTTAAAAATCAAAAAA	1991	
QY	5337	TTGGCGGAGATGTTGGCGGAGCACTGTATTGACGTAATCTGGGAGGCTGAGGCGAGAG	5396	
Db	1992	TTGGCGGAGATGTTGGCGGAGCACTGTATTGACGTAATCTGGGAGGCTGAGGCGAGAG	2051	
QY	5397	AATGGCTGAACCCGAGAGGACAGATTGTCAGTGAAGCCGAGATCGGCCACTGCACTCA	5456	
Db	2052	AATGGCTGAACCCGAGAGGACAGATTGTCAGTGAAGCCGAGATCGGCCACTGCACTCA	2111	
QY	5457	GCCTGGGCGACAGCAAGCTCATCTGTGAAA-AGAAAAAAGAAAGCTTCAAGGTGAG	5515	

Db 2112 GCCTGGGCGACAGCAAGCAAGCTCCATCTGGAAAGAAAAGAAAGCTTCAGGCTGAG 2171
QY 5516 CCAGAGCCCGAGGCTGTAATTCCTGACTACATGACCTTGGCAAGGACCTTCCTCC 5575
Db 2172 CCAGAGCCCGAGGCTGTAATTCCTGACTACATGACCTTGGCAAGGACCTTCCTCC 2231
QY 5576 CTGGCCGAGTTACAGGGGTTGGATGACATCCAGAGTCCCTCCAGATTTACGCTGCAT 5635
Db 2232 CTGGCCGAGTTACAGGGGTTGGATGACATCCAGAGTCCCTCCAGATTTACGCTGCAT 2291
QY 5636 GGTTCCT-AGATGAGAGATGGGAGCTTCCCTCTCTACACCCAGCCGCTGTCCATT 5694
Db 2292 GGTTCCTNAGATGAGAGATGGGAGCTTCCCTCTCTACACCCAGCCGCTGTCCATT 2351
QY 5695 CAGGTGATGACAGGAGGAGTCACTGCTCCCAATCCCGAGTTCCAAAGCCCTTGGG 5752
Db 2352 CAGGTGATGACAGGAGGAGTCACTGCTCCCAATCCCGAGTTCCAAAGCCCTTGGG 2409

RESULT 9
US-09-760-463-27

Sequence 27, Application US/09760463
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT252
CURRENT APPLICATION NUMBER: US/09/760,463
CURRENT FILING DATE: 2001-01-16
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 2409
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (3)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (15)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (56)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (189)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1478)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1465)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1548)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1602)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1603)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1818)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1819)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1824)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

LOCATION: (1828)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1838)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1841)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1875)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (2145)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (2298)
OTHER INFORMATION: n equals a,t,g, or c
US-09-760-463-27

Query Match 35.7%; Score 2225.8; DB 33; Length 2409;
Best Local Similarity 98.7%; Pred. No. 1.6e-257;
Matches 2367; Conservative 0; Mismatches 14; Indels 17; Gaps 12;

QY 3371 CCCTGTCAAACAGACACCTGCGCTCTACATCAGCAGATG-TGGGTGGGCGCAGATA 3429
Db 13 CCCTGTCAAACAGACACCTGCGCTCTACATCAGCAGATG-TGGGTGGGCGCAGATA 72
QY 3430 AGAGATTAAGAGAGGCTGCCGAGCCAGCAGTGGCAACGGCAGAGTCCCTATCCACA 3489
Db 73 AGAGATTAAGAGAGGCTGCCGAGCCAGCAGTGGCAACGGCAGAGTCCCTATCCACA 132
QY 3490 ATATGGCAGCTTTGTTCTTTGCTTGTGCTATTAATCTTCTACTGCTGCTTTTGG 3548
Db 133 ATATGGCAGCTTTGTTCTTTGCTTGTGCTATTAATCTTCTACTGCTGCTTTTGG 192
QY 3549 GTCCACACTGCTTTTATAGCTGTAACTACATCAGCAGAGTCTGACACTTCTGTA 3608
Db 193 GTCCACACTGCTTTTATAGCTGTAACTACATCAGCAGAGTCTGACACTTCTGTA 252
QY 3609 AGCCACTAAGACACAGAGCCACCGGGAGATGAACACTCCGGCGGCTGCTTAAG 3668
Db 253 AGCCACTAAGACACAGAGCCACCGGGAGATGAACACTCCGGCGGCTGCTTAAG 312
QY 3669 AGCTATTAACCTACCGGAGGCTGTGAGTTCACCTCCAGCAGGAGACACAGAC 3728
Db 313 AGCTATTAACCTACCGGAGGCTGTGAGTTCACCTCCAGCAGGAGACACAGAC 372
QY 3729 CCACCAGAGAGAGAACTGCGAACACTGTGAACATCAGAGAGAACTCCAGATGC 3788
Db 373 CCACCAGAGAGAGAACTGCGAACACTGTGAACATCAGAGAGAACTCCAGATGC 432
QY 3789 ACCACTTAAGAGCTGTAACTACATCAGTCCGAGGCTCCGGCTCTTCTTGAAGTCA 3848
Db 433 ACCACTTAAGAGCTGTAACTACATCAGTCCGAGGCTCCGGCTCTTCTTGAAGTCA 492
QY 3849 GAGACCAAGACACTACAGGTTTCGACACAGCCAGAGTGTAGATCAGCCGCGAA 3908
Db 493 GAGACCAAGACACTACAGGTTTCGACACAGCCAGAGTGTAGATCAGCCGCGAA 552
QY 3909 CATGATGAATGAGCCCTCTCTGCAAAAAAATTAACAAAATTTGGCGAGCATGCT 3968
Db 553 CATGATGAATGAGCCCTCTCTGCAAAAAAATTAACAAAATTTGGCGAGCATGCT 612
QY 3969 GGTCCGTGCTGTGGTCCAGACTACGGGGAGGCTAAAGTGGAGATCGCTTGAAGCTG 4028
Db 613 GGTCCGTGCTGTGGTCCAGACTACGGGGAGGCTAAAGTGGAGATCGCTTGAAGCTG 672
QY 4029 GGAGGTGAAGCTGAGTGAAGTGTGATTTGACACAGAGCCCTTAAGTGGGGAGAC 4088
Db 673 GGAGGTGAAGCTGAGTGAAGTGTGATTTGACACAGAGCCCTTAAGTGGGGAGAC 732
QY 4089 TGAGACCTGTTTCCCTCCGCAAAAAAATTAAGTGAAGTGAAGTGAAGTGAAGTGA 4188

Db 733 TGAAGCCCTGTTCCCTCGGCAAAAATTTGACAAAAGTGTATTAAGAGCTCCGTGATA 792
 QY 4149 TGGCTAGGGGCACTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCCGAGGCGG 4208
 Db 793 TGGCTAGGGGCACTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCCGAGGCGG 852
 QY 4209 GTTACTTAAGGTAGAGAGTGTGAGACAGCTTGGCCAACTGGAGAAAGCCCATCTCTTC 4268
 Db 853 GTTACTTAAGGTAGAGAGTGTGAGACAGCTTGGCCAACTGGAGAAAGCCCATCTCTTC 912
 QY 4269 TAAATAATACAAATTAAGCCGCTGTGGGGGAGTGGTGAAGCATGCTGTATATCCAGCT 4328
 Db 913 TAAATAATACAAATTAAGCCGCTGTGGGGGAGTGGTGAAGCATGCTGTATATCCAGCT 972
 QY 4329 ACTCAGAGGCTGAGGCAAGAGATCATCTTGAACCCAGAGAGCGCGGCTTGACAGTGAAG 4388
 Db 973 ACTCAGAGGCTGAGGCAAGAGATCATCTTGAACCCAGAGAGCGCGGCTTGACAGTGAAG 1032
 QY 4389 GAGATGTGCTATGACATCCAGCACTCCAGCTGGGCAACAAGAGCCAACTGTCTCT 4448
 Db 1033 GAGATGTGCTATGACATCCAGCACTCCAGCTGGGCAACAAGAGCCAACTGTCTCT 1092
 QY 4449 TAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4508
 Db 1093 T-AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
 QY 4509 CAACATGTTTAAGATGTGAGAGCTCTGCTTCATGTGCTCTGTATAAAAACCACTCTCA 4568
 Db 1152 CAACATGTTTAAGATGTGAGAGCTCTGCTTCATGTGCTCTGTATAAAAACCACTCTCA 1211
 QY 4569 AGGCGAGGTGCTGCTGCTATATATCCAGCACTTTGGGAGGCGGAGGCGGTG 4628
 Db 1212 AGGCGAGGTGCTGCTGCTATATATCCAGCACTTTGGGAGGCGGAGGCGGTG 1271
 QY 4629 ATCAGCTGAGGTGAGAGTTCGAGACAGCTGACCAACAATGTGGAATCCACCTC 4688
 Db 1272 ATCAGCTGAGGTGAGAGTTCGAGACAGCTGACCAACAATGTGGAATCCACCTC 1331
 QY 4689 TACTAAAAATACAAATTAAGATGAGATGAGTGTGATGCTGTATATCCACCTACTTGG 4748
 Db 1332 TACTAAAAATACAAATTAAGATGAGATGAGTGTGATGCTGTATATCCACCTACTTGG 1391
 QY 4749 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4808
 Db 1392 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
 QY 4809 GTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4866
 Db 1452 GTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1511
 QY 4867 AAAAACCACCTCTCTACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4925
 Db 1512 AAAAACCACCTCTCTACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1571
 QY 4926 GTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4983
 Db 1572 GTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1631
 QY 4984 GAGATGAGTGTGAGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 5043
 Db 1632 GAGATGAGTGTGAGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1691
 QY 5044 ACCAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5103
 Db 1692 ACCAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1751
 QY 5104 GGCAGGCTGGGTTAATTAAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 5163
 Db 1752 GGCAGGCTGGGTTAATTAAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 1811
 QY 5164 CCCAGGCA---GGCTGTGGGCTGT---AGAAAGCTTCAGGCGTAAAGCGGCGGAGG 5217
 Db 1812 CCCAGGNNCAGGNCCTNTGTGGCTGTNTANGAANAACGTTCAAGGCTTAAGGCGGAGGCGG 1871

QY 5218 GGC-TCAGGCTGTATATCCAGGACCTTTGGGAGGCGGAGGCGGGTGTATCAGAGTCAAG 5276
 Db 1872 GGCNTCAGGCTGTATATCCAGGACCTTTGGGAGGCGGAGGCGGGTGTATCAGAGTCAAG 1931
 QY 5277 GAGATGTGACATCTCTGGGTAAACAGAGGTAAACCCCGTCTACTTAAAAATTCAAAAA 5336
 Db 1932 GAGATGTGACATCTCTGGGTAAACAGAGGTAAACCCCGTCTACTTAAAAATTCAAAAA 1991
 QY 5337 TTGGCCGGGATGATGGGCGGCGACCTGTATGTTCCAGTACTCTGAGAGGCTGAGAGGAG 5396
 Db 1992 TTGGCCGGGATGATGGGCGGCGACCTGTATGTTCCAGTACTCTGAGAGGCTGAGAGGAG 2051
 QY 5397 AATGGCTGAACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5456
 Db 2052 AATGGCTGAACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2111
 QY 5457 GCCTGGGCGACAGAGCAAGCTCATCTGGAAT-AGAAAAAGAAAGCTTCAGTGTGAG 5515
 Db 2112 GCCTGGGCGACAGAGCAAGCTCATCTGGAATAGAAAGAAAGAAAGAAAGCTTCAGTGTGAG 2171
 QY 5516 CCAGAGGCGGAGGCTGTATATTCATGCTATACCATGACCTTTGGGCAAGGACCTTCTCC 5575
 Db 2172 CCAGAGGCGGAGGCTGTATATTCATGCTATACCATGACCTTTGGGCAAGGACCTTCTCC 2231
 QY 5576 CTGGCCAGTTTCACGGGCTTGAATGACTCCAAAGTCCCTTCAGCATTTAACGCTGAT 5635
 Db 2232 CTGGCCAGTTTCACGGGCTTGAATGACTCCAAAGTCCCTTCAGCATTTAACGCTGAT 2291
 QY 5636 GGTTC- AAGATGAGAAAGTGGGCAAGTTCCCTCTCATACCCAGGCGGCTGTCCACTT 5694
 Db 2292 GGTTC- AAGATGAGAAAGTGGGCAAGTTCCCTCTCATACCCAGGCGGCTGTCCACTT 2351
 QY 5695 CAAGGTGAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5752
 Db 2352 CAAGGTGAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2409

RESULT 10
 US-09-760-485-1047
 ? Sequence 1047, Application US/09760485
 ? GENERAL INFORMATION:
 ? APPLICANT: Rosen et al.
 ? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ? FILE REFERENCE: P217
 ? CURRENT APPLICATION NUMBER: US/09/760,485
 ? CURRENT FILING DATE: 2001-01-16
 ? Prior application data removed - consult PALM or file wrapper
 ? NUMBER OF SEQ ID NOS: 1477
 ? SOFTWARE: PatentIn Ver. 2.0
 ? SEQ ID NO 1047
 ? LENGTH: 2409
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? FEATURE:
 ? NAME/KEY: SITE
 ? LOCATION: (3)
 ? OTHER INFORMATION: n equals a,t,g, or c
 ? NAME/KEY: SITE
 ? LOCATION: (15)
 ? OTHER INFORMATION: n equals a,t,g, or c
 ? NAME/KEY: SITE
 ? LOCATION: (56)
 ? OTHER INFORMATION: n equals a,t,g, or c
 ? NAME/KEY: SITE
 ? LOCATION: (189)
 ? OTHER INFORMATION: n equals a,t,g, or c
 ? NAME/KEY: SITE
 ? LOCATION: (1465)
 ? OTHER INFORMATION: n equals a,t,g, or c
 ? NAME/KEY: SITE
 ? LOCATION: (1478)
 ? OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1548)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1602)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1603)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1818)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1824)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1828)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1838)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1841)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1875)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (2145)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (2298)
OTHER INFORMATION: n equals a,t,g, or c
US-09-760-485-1047

Query Match 35.7%; Score 2225.8; DB 33; Length 2409;
Best Local Similarity 98.7%; Pred. No. 1.6e-257;
Matches 2367; Conservative 0; Mismatches 14; Indels 17; Gaps 12;

3371 CCCTGTCAAAACAGACACTGGCTCTACCAATTCAGACAGATG-TGGGTGGGGCCAGATA 3429
13 CCNTGTCAAAACAGACACTGGCTCTACCAATTCAGACAGATG-TGGGTGGGGCCAGATA 72
3430 AGAGATAAAACAGACACTGGCTCTACCAATTCAGACAGATG-TGGGTGGGGCCAGATA 3489
73 AGAGATAAAACAGACACTGGCTCTACCAATTCAGACAGATG-TGGGTGGGGCCAGATA 132
3430 ATATGGCAGCTTGTCTTTTGTCTTTGCGATTAATCTTGTCTACTGCTCGCTTTT-TGG 3548
133 ATATGGCAGCTTGTCTTTTGTCTTTGCGATTAATCTTGTCTACTGCTCGCTTTTNTGG 192
3549 GTCCACACTGCTTTATGAGCTGTAAACATCAGACAGAAAGTCTGAGCTTCACTCCTGA 3608
193 GTCCACACTGCTTTATGAGCTGTAAACATCAGACAGAAAGTCTGAGCTTCACTCCTGA 252
3609 AGCCACTAAGACAGACAGCCACCGGAGGAATGAACAATCCGGCCGCGCTTGAAG 3668
253 AGCCACTAAGACAGACAGCCACCGGAGGAATGAACAATCCGGCCGCGCTTGAAG 312
3669 AGCTATAACACTCAGCGGGAAGTGTGACGCTTCTCTCAGCCAGACAGACAGAAC 3728
313 AGCTATAACACTCAGCGGGAAGTGTGACGCTTCTCTCAGCCAGACAGACAGAAC 372
3729 CCACCAAGAGAAAGAACTGCGAACAATCTGAACATCAGAAAGAAACAATCTCAGATG 3788
373 CCACCAAGAGAAAGAACTGCGAACAATCTGAACATCAGAAAGAAACAATCTCAGATG 432
3789 ACCACCTTAAGAGCTGTAACTCACTGCGAGGGTCCGGGCTTCTCTTGAAGTCACT 3848
433 ACCACCTTAAGAGCTGTAACTCACTGCGAGGGTCCGGGCTTCTCTTGAAGTCACT 492

3849 GAGACCAACCACTACACAGTTTGGACACCAAGCCAGAGTTTGAATCAGCTGGGCAA 3908
493 GAGACCAACCACTACACAGTTTGGACACCAAGCCAGAGTTTGAATCAGCTGGGCAA 552
3909 CATGATGAATGCTCTCTCTGCAAAAAAATTTACAAAATTTGGCGGAGCATGCT 3968
553 CATGATGAATGCTCTCTCTGCAAAAAAATTTACAAAATTTGGCGGAGCATGCT 612
3969 GGTCCGCTGCTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGAGTCTTGAAGCTG 4028
613 GGTCCGCTGCTGTGTCCAGCTACGCGGAGGCTAAAGTGGAGAGTCTTGAAGCTG 672
4029 GGAGGTGAAGACTCAGATGAGTGTATGTAACACAGCCCTCTAGGCTGGGGGACAGC 4088
673 GGAGGTGAAGACTCAGATGAGTGTATGTAACACAGCCCTCTAGGCTGGGGGACAGC 732
4089 TGAGACCCCTGTTCCCTCCGCAAAAAAATTTACAAAATTTGAAGAGTGCCTGATA 4148
733 TGAGACCCCTGTTCCCTCCGCAAAAAAATTTACAAAATTTGAAGAGTGCCTGATA 792
4149 TGGCTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGAAAGCCGAGCGGCG 4208
793 TGGCTAGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGAAAGCCGAGCGGCG 852
4209 GTACACTAAGTCAAGAGTGTGTAACCAAGCCCTGGCCAACTGAGAAAGCCCATCTTC 4268
853 GTACACTAAGTCAAGAGTGTGTAACCAAGCCCTGGCCAACTGAGAAAGCCCATCTTC 912
4269 TAAAAATACAAATTTAGCCGCTGTGGGGGAGTGTGAGCATGCTTAATCCAGCT 4328
913 TAAAAATACAAATTTAGCCGCTGTGGGGGAGTGTGAGCATGCTTAATCCAGCT 972
4329 ACTCAGAGAGCTGAGGCGAGAGATCATTGAACCCAGAGCGCGGTTGCAATGAGCC 4388
973 ACTCAGAGAGCTGAGGCGAGAGATCATTGAACCCAGAGCGCGGTTGCAATGAGCC 1032
4389 GAGATGCTGCCATGTGACCTCCACCACTCCAGCTGGGGCAAGAGCCAACTGCTTC 4448
1033 GAGATGCTGCCATGTGACCTCCACCACTCCAGCTGGGGCAAGAGCCAACTGCTTC 1092
4449 TAAAAAAGTGGCTGACATATTAAGAGTGTGCAATGCAATAGTTGGCAG 4508
1093 T-AAAAAAGTGGCTGACATATTAAGAGTGTGCAATGCAATAGTTGGCAG 1151
4509 CAACATGTTTAAGATGTGAGCTCTGCTTCATGCTGTGTTAAAAACCACTCTCA 4568
1152 CAACATGTTTAAGATGTGAGCTCTGCTTCATGCTGTGTTAAAAACCACTCTCA 1211
4569 AGGCCAGTGCAGTGGCTCATGCTTAATCCAGCACTTTGGAGAGCCGAGCGGCTG 4628
1212 AGGCCAGTGCAGTGGCTCATGCTTAATCCAGCACTTTGGAGAGCCGAGCGGCTG 1271
4629 ATACCTGAGGTGAGAGTTCGAGACAGCTGACCAACCAATAGTGAATCCACCTC 4688
1272 ATACCTGAGGTGAGAGTTCGAGACAGCTGACCAACCAATAGTGAATCCACCTC 1331
4689 TACTAAAAATACAAAATTTAGATGAGCATGCTGTGATCCTGTAATCCACTACTTGG 4748
1332 TACTAAAAATACAAAATTTAGATGAGCATGCTGTGATCCTGTAATCCACTACTTGG 1391
4749 GAGGCTGAGGCAAGAAATCACTAGAACAGGAGGCGGAGGTTGAGTGAAGCCGAGTC 4808
1392 GAGGCTGAGGCAAGAAATCACTAGAACAGGAGGCGGAGGTTGAGTGAAGCCGAGTC 1451
4809 GTGCAATTCAGCT-CCAGCCTGAGCA-ATGAGCGAAATCTCAATCTCAAAAAAACAAC 4866
1452 GTGCAATTCAGCTNCCAGCCTGAGCAATGAGGAGAAATCTCAATCTCAAAAAAACAAC 1511
4867 AAAAACCACTCTTACTCTCCAGAGGAGTGGGTACA-GAGCTGGGCAATCAGTCAAG 4925
1512 AAAAACCACTCTTACTCTCCAGAGGAGTGGGTACANAGAGTGGGCAATCAGTCAAG 1571

QY 4926 GTGCTGAGCCACAGAGCTAAGCGGAGCTG--CAGAGCCGGGACAGATAACAGTGTGT 4983
Db 1572 GTGCTGAGCCACAGAGCTAAGCGGAGCTG--CAGAGCCGGGACAGATAACAGTGTGT 1631
QY 4984 GAGATCACTGTGTGATGATCAGACGCTCCCTGCAATTTGGTACACACAGGGGCCCCCAAGC 5043
Db 1632 GAGATCACTGTGTGATGATCAGACGCTCCCTGCAATTTGGTACACACAGGGGCCCCCAAGC 1691
QY 5044 ACCAGAGATGAGCCCATCCAGTCAACACATCCACTTCCATCCAGAGATGCTGTCTT 5103
Db 1692 ACCAGAGATGAGCCCATCCAGTCAACACATCCACTTCCATCCAGAGATGCTGTCTT 1751
QY 5104 GGCACGCTGGGGTAATTAAGACAGAAAGTGACAGTCTTGGGTGTGTGCTACAGACTGC 5163
Db 1752 GGCACGCTGGGGTAATTAAGACAGAAAGTGACAGTCTTGGGTGTGTGCTACAGACTGC 1811
QY 5164 CCCAGGCA---GGCCTGTGTGCTGT--AGAAAAGCTTACAGGCGCTAGCCGGGACAGGT 5217
Db 1812 CCCAGGNNACAGGCTTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1871
QY 5218 GGC--TCAAGCCTGTATATCCAGCACTTTGGAGGCGGAGGGGGGTGTGTGTGTGTGTGT 5276
Db 1872 GGCNTCAAGCCTGTATATCCAGCACTTTGGAGGCGGAGGGGGGTGTGTGTGTGTGTGTGT 1931
QY 5277 GAGATCGTACCATCTGT 5336
Db 1932 GAGATCGTACCATCTGT 1991
QY 5337 TTGGCCGGGCAATGT 5396
Db 1992 TTGGCCGGGCAATGT 2051
QY 5397 AATGCGGTGAACCCGAGAGGAGAGTTTGCAGTACGAGAGATGCGGCACTGTGTGTGTGT 5456
Db 2052 AATGCGGTGAACCCGAGAGGAGAGTTTGCAGTACGAGAGATGCGGCACTGTGTGTGT 2111
QY 5457 GCCTGGGCGACAGAGCAAGACTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5515
Db 2112 GCCTGGGCGACAGAGCAAGACTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2171
QY 5516 CCAGAGGCGCCAGGCTGTATATCTGTCACTTACATGACCTTGGGCAAGGCACTGTGTGT 5575
Db 2172 CCAGAGGCGCCAGGCTGTATATCTGTCACTTACATGACCTTGGGCAAGGCACTGTGTGT 2231
QY 5576 CTGGCCCAAGTTACAGGGGTTGGAATCGACTCCAGAGTCCCTTCAGCATTAAGCTGCAT 5635
Db 2232 CTGGCCCAAGTTACAGGGGTTGGAATCGACTCCAGAGTCCCTTCAGCATTAAGCTGCAT 2291
QY 5636 GGTTCCT--AAGATGAGAAAGATGGGCGAGTTCCCTCTCTCAGACCCAGCCGCTGTGCATT 5694
Db 2292 GGTTCCTMAAGATGAGAAAGATGGGCGAGTTCCCTCTCTCAGACCCAGCCGCTGTGCATT 2351
QY 5695 CAAGGTGAATGACAGGGAAGTCAACGCTGCCAATCCCGAGTTCCAAAGCCCTTGGG 5752
Db 2352 CAAGGTGAATGACAGGGAAGTCAACGCTGCCAATCCCGAGTTCCAAAGCCCTTGGG 2409

RESULT 11

US-10-143-877-27
; Sequence 27 Application US/10143877
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT252CIN
; CURRENT APPLICATION NUMBER: US/10/143,877
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (56)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1465)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1478)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1602)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1603)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1818)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1819)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1824)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1828)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1838)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1841)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1875)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2145)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2298)
OTHER INFORMATION: n equals a,t,g, or c
US-10-143-877-27

Query Match 35.7%; Score 2225.8; DB 46; Length 2409;
Best Local Similarity 98.7%; Pred. No. 1.6e-257;
Matches 2367; Conservative 0; Mismatches 14; Indels 17; Gaps 12;

QY 3371 CCCTGTCAAAACAGACACTGGGCTCTACCAATCAGCAGATG--TGGGTGGGGCCAGATA 3429
DB 13 CCNMTCAAAACAGACACTGGGCTCTACCAATCAGCAGATGNTGGGTGGGGCCAGATA 72
QY 3430 AGAGATTAAGACAGAGCTGCCGAGCAGAGTGGCAAGCGCACAGGTCCCTATCCACA 3489
DB 73 AGAGATTAAGACAGAGCTGCCGAGCAGAGTGGCAAGCGCACAGGTCCCTATCCACA 132
QY 3490 AATAGGCACTTTGTTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTG 3548
DB 133 AATAGGCACTTTGTTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTG 192
QY 3549 GTCCACACTGCTTTTATGAGCTGTACACTACACAGAGGCTCTGACCTTCACTCTGA 3608
DB 193 GTCCACACTGCTTTTATGAGCTGTACACTACACAGAGGCTCTGACCTTCACTCTGA 252
QY 3609 AGCCACTTAAGACAGAGCCCAACGGGAGAGTAACAACTCCGGCCGCTGCTTAAAG 3668
DB 253 AGCCACTTAAGACAGAGCCCAACGGGAGAGTAACAACTCCGGCCGCTGCTTAAAG 312
QY 3669 AGCTATACACTACACCGGAGAGTCTGAGCTTCACTCTCTCAAGCCAGAGACACGAAAC 3728
DB 313 AGCTATACACTACACCGGAGAGTCTGAGCTTCACTCTCTCAAGCCAGAGACACGAAAC 372
QY 3729 CCACAGAGAGAGAAATCGGAAACATCTGAACATCAGAAAGAAACAACTCCAGATGC 3788
DB 373 CCACAGAGAGAGAAATCGGAAACATCTGAACATCAGAAAGAAACAACTCCAGATGC 432
QY 3789 ACCACCTTAAGAGCTGTACACTCACTCGAGAGGTCCGGCTCTCTTTTGAATCACT 3848
DB 433 ACCACCTTAAGAGCTGTACACTCACTCGAGAGGTCCGGCTCTCTTTTGAATCACT 492
QY 3849 GAGACCAAGCACTACACGTTTGGGACAAAGCCAGAGATTTGATCAGCTGGGAA 3908
DB 493 GAGACCAAGCACTACACGTTTGGGACAAAGCCAGAGATTTGATCAGCTGGGAA 552
QY 3909 CATGATGAATGCCCCCTCTCTGCAAAAAAAATTTACAAAAATTTGGCGAGCATGCT 3968
DB 553 CATGATGAATGCCCCCTCTCTGCAAAAAAAATTTACAAAAATTTGGCGAGCATGCT 612
QY 3969 GGTCCGTGCTGTGTGCTCCAGCTACAGCGGAGGCTAAAGTGGAGATCGCTTGAGCTG 4028
DB 613 GGTCCGTGCTGTGTGCTCCAGCTACAGCGGAGGCTAAAGTGGAGATCGCTTGAGCTG 672
QY 4029 GGAGGTGAAGAGTGCAGTGTGATTTGATACACAGCCCTCTAGGCTGGGGGACAGAC 4088
DB 673 GGAGGTGAAGAGTGCAGTGTGATTTGATACACAGCCCTCTAGGCTGGGGGACAGAC 732
QY 4089 TGAGACCTGTTTCCCTCCGCAAAAAAAATTTGACAAAAAGTGAATAGAGTGCCTGATA 4148
DB 733 TGAGACCTGTTTCCCTCCGCAAAAAAAATTTGACAAAAAGTGAATAGAGTGCCTGATA 792
QY 4149 TGGCTAGGCGCAGTGTGCTGATCTGATATCCAGCACTTTGGAGCCGAGGCGGG 4208
DB 793 TGGCTAGGCGCAGTGTGCTGATCTGATATCCAGCACTTTGGAGCCGAGGCGGG 852
QY 4209 GTACACCTAAGGTCAGAGTGTGAGACAGCCGTCGCAACATGAGAGAAAGCCATCTTC 4268
DB 853 GTACACCTAAGGTCAGAGTGTGAGACAGCCGTCGCAACATGAGAGAAAGCCATCTTC 912
QY 4269 TAAAAATACAAATTAAGCGGCTGTGGGGCAGTGGTGGAGCATGCTGTATCCAGCT 4328
DB 913 TAAAAATACAAATTAAGCGGCTGTGGGGCAGTGGTGGAGCATGCTGTATCCAGCT 972
QY 4329 ACTCAGAGGCTGAGGCGAGAGATCACTTGAACCCAGAGAGCGGCTTGACGTGAGCC 4388
DB 973 ACTCAGAGGCTGAGGCGAGAGATCACTTGAACCCAGAGAGCGGCTTGACGTGAGCC 1032

QY 4389 GAGATGTCGTCACATGCTCCACCCATCTCAGCTGTGGCAACAGAGCCAACTGCTCT 4448
DB 1033 GAGATGTCGTCACATGCTCCACCCATCTCAGCTGTGGCAACAGAGCCAACTGCTCT 1092
QY 4449 TAAAAAAAGTGGCTGACATATTAAGAGTGTGTGCAATGCAATAGTTCAGG 4508
DB 1093 T-TAAAAAAAGTGGCTGACATATTAAGAGTGTGTGCAATGCAATAGTTCAGG 1151
QY 4509 CAACATGTTTAAAGATGTGGAGCTCTGCTTCCATGCTGCTTTAAAAAACCACCTCA 4568
DB 1152 CAACATGTTTAAAGATGTGGAGCTCTGCTTCCATGCTGCTTTAAAAAACCACCTCA 1211
QY 4569 AGGCGAGGTCAGTGTGCTATGCTATATCCACACTTTGGAGGCGGAGGCGGTGG 4628
DB 1212 AGGCGAGGTCAGTGTGCTATGCTATATCCACACTTTGGAGGCGGAGGCGGTGG 1271
QY 4629 ATCACCTGAGTCAGAGTTCAGACAGAGCCGACACCAACATGATGTAATCCACCTC 4688
DB 1272 ATCACCTGAGTCAGAGTTCAGACAGAGCCGACACCAACATGATGTAATCCACCTC 1331
QY 4689 TACTTAAATATCAAAATTAAGATGAGCATGTGTGCTGATCTGATCCACCTACTGG 4748
DB 1332 TACTTAAATATCAAAATTAAGATGAGCATGTGTGCTGATCTGATCCACCTACTGG 1391
QY 4749 GAGGCTGAGGCGAGAAATCACTAGAACAGGAGGCGGAGGTTGTAGTGAAGCGAGATC 4808
DB 1392 GAGGCTGAGGCGAGAAATCACTAGAACAGGAGGCGGAGGTTGTAGTGAAGCGAGATC 1451
QY 4809 GTGCCATGTCACCT--CCAGCTGAGCA--ATGAGCGAACTCCATCTCAAAAAAACAAC 4866
DB 1452 GTGCCATGTCACCTNCCAGCTGAGCAANTGAGCGAACTCCATCTCAAAAAAACAAC 1511
QY 4867 AAAAAACAATCTTACTTCCAGGAGCTGGGTACA--GAGTGGCCACATCAGTGCAG 4925
DB 1512 AAAAAACAATCTTACTTCCAGGAGCTGGGTACANAGAGTGGGCCATCAGTGCAG 1571
QY 4926 GTGTCAGGACAGAGCTTAAGCGGAGGCTG--CAGGACCGGAGCAGATACAGTGT 4983
DB 1572 GTGTCAGGACAGAGCTTAAGCGGAGGCTG--CAGGACCGGAGCAGATACAGTGT 1631
QY 4984 GAGATCACTGTGTGATGATCAGACGTCCTGCTGATGTGACACCAAGGCGGCCCAAGC 5043
DB 1632 GAGATCACTGTGTGATGATCAGACGTCCTGCTGATGTGACACCAAGGCGGCCCAAGC 1691
QY 5044 ACCAGAGATGGCCCATCCAGTCACCAATCCTTCTCATCAGAGATGTGTTTCT 5103
DB 1692 ACCAGAGATGGCCCATCCAGTCACCAATCCTTCTCATCAGAGATGTGTTTCT 1751
QY 5104 GGCACGTCGGGTAATTAAGACAGAGGTGACACTTGGGTGTGTGTCAGTCAAGCTGC 5163
DB 1752 GGCACGTCGGGTAATTAAGACAGAGGTGACACTTGGGTGTGTGTCAGTCAAGCTGC 1811
QY 5164 CCCAGGCA---GGCCTGTGTGCTGT--AGAAAACTTCAGGCTAGGCGGCGACGCT 5217
DB 1812 CCCAGGNNCAGGNCCTTNTGTGGCTGTGNTANGAAAACTTCAGGCTAGGCGGCGACGCT 1871
QY 5218 GGC-TGACGCTGTATCCAGCACTTTGGAGGCGGAGGCGGTGTGATCAGAGCTCAG 5276
DB 1872 GGCNTCAGGCTGTATCCAGCACTTTGGAGGCGGAGGCGGTGTGATCAGAGCTCAG 1931
QY 5277 GAGATGTCGACCATCTGCTGCTTAACAGGCGGAAACCCGCTCTACTAAAAATACAAAA 5336
DB 1932 GAGATGTCGACCATCTGCTGCTTAACAGGCGGTAACCCCTCTCTACTAAAAATACAAAA 1991
QY 5337 TTGGCGGCGCATGTGGCGGCGACCTGTATGTTCCAGTACTCTGGAGGCTGAGGAGAG 5396
DB 1992 TTGGCGGCGCATGTGGCGGCGACCTGTATGTTCCAGTACTCTGGAGGCTGAGGAGAG 2051
QY 5397 AATGCGTGTAAACCCAGAGGCGAGAGTGTGTCAGTGTGAGTGTGCGCCACTGCACTCA 5456
DB 2052 AATGCGTGTAAACCCAGAGGCGAGAGTGTGTCAGTGTGAGTGTGCGCCACTGCACTCA 2111
QY 5457 GCCTGGGCGACAGCAAGACTCCATCTGGAAT--AGAAAAAGAAAACTTCAGTGTGAG 5515

Db 2112 GCGTGGGCGACAGAGACGCTCATCTGGAANAGAAAAAGAAAGCTTCAGGCTGAG 2171
QY 5516 CCAGAGGCCAGGCTGTAATTCCTGTCACCTTACCATACCTTGGGCAAGGACCTTCC 5575
Db 2172 CCAGAGGCCAGGCTGTAATTCCTGTCACCTTACCATACCTTGGGCAAGGACCTTCC 2231
QY 5576 CTGGCCAGCTTACGCGGGTTGGAATGACTCCAGGTCCTTCACGATTAAACGTCGAT 5635
Db 2232 CTGGCCAGCTTACGCGGGTTGGAATGACTCCAGGTCCTTCACGATTAAACGTCGAT 2291
QY 5636 GGTTCCT-AAGATGAGAAGATGGGCGATTCCCTCTCCACCCGAGCCGCTGCACTT 5694
Db 2292 GGTTCCTAAAGATGAGAAGATGGGCGATTCCCTCTCCACCCGAGCCGCTGCACTT 2351
QY 5695 CAAGGTGAATGACAGGAGATGACGTCGTCACCTCCGAGGATTCGCAAAAGCCCTTGGG 5752
Db 2352 CAAGGTGAATGACAGGAGATGACGTCGTCACCTCCGAGGATTCGCAAAAGCCCTTGGG 2409

RESULT 12
US-10-216-436-1047
; Sequence 1047, Application US/10216436
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antihodies
; FILE REFERENCE: P0217CIN
; CURRENT APPLICATION NUMBER: US/10/216,436
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/760,485
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1477
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (56)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (189)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:

NAME/KEY: misc_feature
LOCATION: (1465)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1478)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1548)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1602)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1603)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1818)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1819)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1824)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1828)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1838)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1841)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1875)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2145)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2298)
OTHER INFORMATION: n equals a,t,g, or c
US-10-216-436-1047

Query Match 35.7%; Score 2225.8; DB 48; Length 2409;
Best Local Similarity 98.7%; Pred. No. 1.6e-257;
Matches 2367; Conservative 0; Mismatches 14; Indels 17; Gaps 12;

QY 3371 CCCTGTCAAAACAGACACCTCGGCTTACCAATCAGCAGATG-TGGGTGGGCGCAGATA 3429
Db 13 CCNTGTCAAAACAGACACCTCGGCTTACCAATCAGCAGATGTTGGGGCGCAGATA 72
QY 3430 AGAGAAATTAAGAGGCTGCGGAGCAGAGGCAAGCGCAGAGGTCCTATCCACA 3489
Db 73 AGAGAAATTAAGAGGCTGCGGAGCAGAGGCAAGCGCAGAGGTCCTATCCACA 132
QY 3490 ATATGCGAGCTTGTCTTTGCTGTGTGCGATAAATCTTGCTACTGCTGCTTTT-TGG 3548
Db 133 ATATGCGAGCTTGTCTTTGCTGTGTGCGATAAATCTTGCTACTGCTGCTTTTGG 192

QY	3549	GTCCACACTGCTTTTATAGAGCTGTAACTACCTACCAAGAGGTCTGACGCTTCACTCTGA	3608
Db	193	GTCCACACTGCTTTTATAGAGCTGTAACTACCTACCAAGAGGTCTGACGCTTCACTCTGA	252
QY	3609	AGCCACTAAGACCAACGAGCCCAACGGGGGAATGAACAATCCCGCCGCGCTTGCTTAAG	3668
Db	253	AGCCACTAAGACCAACGAGCCCAACGGGGGAATGAACAATCCCGCCGCGCTTGCTTAAG	312
QY	3669	AGCTATTAACATCACTACCGCGAAGGTCTGAGCTTCACTCTCAGCAGAGGAACCAACGAA	3728
Db	313	AGCTATTAACATCACTACCGCGAAGGTCTGAGCTTCACTCTCAGCAGAGGAACCAACGAA	372
QY	3729	CCACCAGAAGGAAGAACTGCGAAACATCTGAACATCAGAAGGAACAACTCCAGATGC	3788
Db	373	CCACCAGAAGGAAGAACTGCGAAACATCTGAACATCAGAAGGAACAACTCCAGATGC	432
QY	3789	ACCACTTAAAGAGCTGTAACTACCTACCTACGAGGGTCGCGCGCTTCTTTGAAGTACGT	3848
Db	433	ACCACTTAAAGAGCTGTAACTACCTACCTACGAGGGTCGCGCGCTTCTTTGAAGTACGT	492
QY	3849	GAGACCAAGCACTACCAAGTTTGGAGACAAGCCAGAGAGTTTGAGATCAGCTGGCGAA	3908
Db	493	GAGACCAAGCACTACCAAGTTTGGAGACAAGCCAGAGAGTTTGAGATCAGCTGGCGAA	552
QY	3909	CATGATGAATAGCCCTCTCTGCAAAAAAATTTACAAAAATTTGCGCGAGCATGT	3968
Db	553	CATGATGAATAGCCCTCTCTGCAAAAAAATTTACAAAAATTTGCGCGAGCATGT	612
QY	3969	GGTCCGTCGCTCTGTGTGCCAGCTACGCGGGAGGCTAAAGTGGAGAGATCGTTAGCCTG	4028
Db	613	GGTCCGTCGCTCTGTGTGCCAGCTACGCGGGAGGCTAAAGTGGAGAGATCGTTAGCCTG	672
QY	4029	GGAGGTGAAGAGCTCAGTGAAGTGTAAACCAAGCCCTTGAAGCTGGGGGAACAGAC	4088
Db	673	GGAGGTGAAGAGCTCAGTGAAGTGTAAACCAAGCCCTTGAAGCTGGGGGAACAGAC	732
QY	4089	TGAGACCCCTGTTCCCTCCGCAAAAAAATTTAGCAAAAAATTAAGAGTGGCTGTATA	4148
Db	733	TGAGACCCCTGTTCCCTCCGCAAAAAAATTTAGCAAAAAATTAAGAGTGGCTGTATA	792
QY	4149	TGGCTAGGGCGAGTGGCTCATAGCTGTAAATCCCAAGACTTTGGGAAGCCGAGCGGGCG	4208
Db	793	TGGCTAGGGCGAGTGGCTCATAGCTGTAAATCCCAAGACTTTGGGAAGCCGAGCGGGCG	852
QY	4209	GTCACCTAAGGCTCAGAGTGTGAGACCAAGCCGGGCAACATGGGAAGGCCATCTCTTC	4268
Db	853	GTCACCTAAGGCTCAGAGTGTGAGACCAAGCCGGGCAACATGGGAAGGCCATCTCTTC	912
QY	4269	TAAAAAATCAAAAATTAAGCCGGCTGTGGGGCGAGTGGTGAACATGAGCCGTAAATCCAGCT	4328
Db	913	TAAAAAATCAAAAATTAAGCCGGCTGTGGGGCGAGTGGTGAACATGAGCCGTAAATCCAGCT	972
QY	4329	ACTCAGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGGTTCAGTGAACC	4388
Db	973	ACTCAGAGAGGCTGAGGCGAGAGAAATCACTTGAACCCAGAGGCGGGTTCAGTGAACC	1032
QY	4389	GAGATCGGCAATGCACTCCACCCACTCCAGCCCTGGGGCAACAAGAGCCAAATCTGTCT	4448
Db	1033	GAGATCGGCAATGCACTCCACCCACTCCAGCCCTGGGGCAACAAGAGCCAAATCTGTCT	1092
QY	4449	TAAAAAATCAAAAATTAAGTGCCTGACATATTAAGAGTGTGCAATGCAATAGTTGGCAGG	4508
Db	1093	T-TAAAAAATCAAAAATTAAGTGCCTGACATATTAAGAGTGTGCAATGCAATAGTTGGCAGG	1151
QY	4509	CAACATGTTTAAAGATGTGAGCTTCCTGCTTCATGGTCTGTTAAAAAACCAACCTCA	4568
Db	1152	CAACATGTTTAAAGATGTGAGCTTCCTGCTTCATGGTCTGTTAAAAAACCAACCTCA	1211
QY	4569	AGGCGAGGTGAGAGGCTCATGCTTAATATCCCAAGCACTTTGGGAAGCCGAGGGGGTGG	4628
Db	1212	AGGCGAGGTGAGAGGCTCATGCTTAATATCCCAAGCACTTTGGGAAGCCGAGGGGGTGG	1271
QY	4629	ATCACCTGAGGTCAAGGATTTGAACCAAGCCTTGACCAACCAATGTTGAATCCACCTC	4688

Db	1272	ATCACCCTGAGGTCAGAGAGTTCGAGACCGACCTGACACAACTGCTAAATCCCACTC	1331
Qy	4689	TACTTAAAAATACAAAAATTAGATGACATAGGTGTCATGCGCTGTAAATCCCACTACTTGG	4748
Db	1332	TACTTAAAAATACAAAAATTAGATGACATAGGTGTCATGCGCTGTAAATCCCACTACTTGG	1391
Qy	4749	GAGCGTGAAGCAGGAAAAATCACTAGAACAGGAGGCGAGGTTGTAGTGAGCCGAGATC	4808
Db	1392	GAGCGTGAAGCAGGAAAAATCACTAGAACAGGAGGCGAGGTTGTAGTGAGCCGAGATC	1451
Qy	4809	GTGCGATTGCACT-CCAGCCTGAGCA-ATGAGCGAAATCCATCTCAAAAAACAACAC	4866
Db	1452	GTGCGATTGCACTTCCAGCGCTGAGCANATGAGCGAAATCCATCTCAAAAAACAACAC	1511
Qy	4867	AAAAAACCACCTTCTACTCTCCAGGAGAGCTGGGTTACA-GAGCTGGGCGCACATGATGCAAG	4925
Db	1512	AAAAAACCACCTTCTACTCTCCAGGAGAGCTGGGTTACANAGCTGGGCGCACATGATGCAAG	1571
Qy	4926	GTGCTGAGGCACAGAGCTAAGCGCGAGCTG--CAGGACCGGGGACCAATTAACAGTGTAT	4988
Db	1572	GTGCTGAGGCACAGAGCTAAGCGCGAGCTGANNACAGAGACCGGAGCAATTAACAGTGTAT	1631
Qy	4984	GAGATCAGTGTGTGATCAGACGTCCTCTGCCATTGTGTGACACCAGGCGGCCCCCAAGC	5043
Db	1632	GAGATCAGTGTGTGATCAGACGTCCTCTGCCATTGTGTGACACCAGGCGGCCCCCAAGC	1691
Qy	5044	ACCAGAGATGCCCCCATCATCAGTCACACCATCATCTTCTCATCAGAGATGTCTGTTCTT	5103
Db	1692	ACCAGAGATGCCCCCATCATCAGTCACACCATCATCTTCTCATCAGAGATGTCTGTTCTT	1751
Qy	5104	GGCACGCTGGGGTTAAATTAGGACAGACAAGGTGCATCTTGGGTGGTGGACGACAGACATGC	5163
Db	1752	GGCACGCTGGGGTTAAATTAGGACAGACAAGGTGCATCTTGGGTGGTGGACGACAGACATGC	1811
Qy	5164	CCCAAGCA---GECCTGTGGCCCTGT--AGAAAACTTCAGCGCTAGCGCGGACGCT	5217
Db	1812	CCCAAGNNCAGGNCCTTGTGGCCCTGTANAGAAAACTTCAGCGCTAGCGCGGACGCT	1871
Qy	5218	GGC-TCAGCCCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGGGGTGCATCAGAGATGAG	5276
Db	1872	GGCNCATCAGCCCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGGGGTGCATCAGAGATGAG	1931
Qy	5277	GAGATCGTAGACCATCTCTGGCTTAACACGGGTGAACCCCGCTCTACTTAAAAATACAAAAA	5338
Db	1932	GAGATCGTAGACCATCTCTGGCTTAACACGGGTGAACCCCGCTCTACTTAAAAATACAAAAA	1991
Qy	5337	TTGGCCGGGCAATGATGGCGGCGACCTGTAGTTCCAGTACTTCGGGAGGCTGAGGCAAGAG	5396
Db	1992	TTGGCCGGGCAATGATGGCGGCGACCTGTAGTTCCAGTACTTCGGGAGGCTGAGGCAAGAG	2051
Qy	5397	AATGGCGTGAACCCGAGAGGCAAGAGATTTGCACTGAGCCGAGATGCGCCACTGCATCCA	5456
Db	2052	AATGGCGTGAACCCGAGAGGCAAGAGATTTGCACTGAGCCGAGATGCGCCACTGCATCCA	2111
Qy	5457	GCGTGGGGACAGAGACCAAGCACTCCATCTGGAAAAAGAAAAAGAAACCTTACAGGCTGAG	5515
Db	2112	GCGTGGGGACAGAGACCAAGCACTCCATCTGGAAAAAGAAAAAGAAACCTTACAGGCTGAG	2171
Qy	5516	CCAGAGGCCACAGCGCTAATCTCTGTCACTTACATGACCTTGGGCAAGGCAACTTCTCTCC	5579
Db	2172	CCAGAGGCCACAGCGCTAATCTCTGTCACTTACATGACCTTGGGCAAGGCAACTTCTCTCC	2231
Qy	5576	CTGGCCCACTTACAGCGGCTTGAATCGACTCCAAAGTCCCTTCCAGCATTAACGCTGCAT	5635
Db	2232	CTGGCCCACTTACAGCGGCTTGAATCGACTCCAAAGTCCCTTCCAGCATTAACGCTGCAT	2291
Qy	5636	GGTTCCT-AAGATGAGAAAGATGGGGCAGTTCCCTCTCTCAACCCAGCCGCGTCCACTT	5694
Db	2292	GGTTCCTAAGATGAGAAAGATGGGGCAGTTCCCTCTCTCAACCCAGCCGCGTCCACTT	2351
Qy	5695	CAAGGTGATATACAGGGAAGTCAAGTTCCTCAATCCCGCAGTTCCAAAGCCCTTGGG	5752

|||||
Db 733 TGAGACCTGTTCCCTCCGCCAAAAATTGACAAAAGTGAATAGAGGTGCTGATA 792
Qy 4149 TGGCTAGGCGGAGGCGTCATGCTGTATATCCAGACATTTTGGGAAGCGGAGCGGG 4208
Db 793 TGGCTAGGCGGAGGCGTCATGCTGTATATCCAGACATTTTGGGAAGCGGAGCGGG 852
Qy 4209 GTCACTTAAGSTCAGAGGTGTAGACACAGCTGGCGCAATGAGAAAGCCATCTTTC 4268
Db 853 GTCACTTAAGSTCAGAGGTGTAGACACAGCTGGCGCAATGAGAAAGCCATCTTTC 912
Qy 4269 TAAAAATACAAATTTAGCCGCTGTGGGGGAGTGTGGAGCATGCTGTATCCAGCT 4328
Db 913 TAAAAATACAAATTTAGCCGCTGTGGGGGAGTGTGGAGCATGCTGTATCCAGCT 972
Qy 4329 ACTGAGGAGGCTGAGGCGGAGGAGATCACTTGAACCCAGAGAGCGCGGTGTCAGTAC 4388
Db 973 ACTGAGGAGGCTGAGGCGGAGGAGATCACTTGAACCCAGAGAGCGCGGTGTCAGTAC 1032
Qy 4389 GAGATCGTGCATTCGACTCCACCCTCCAGCCCTGGGCAACAGAGCCAACTCTGCT 4448
Db 1033 GAGATCGTGCATTCGACTCCACCCTCCAGCCCTGGGCAACAGAGCCAACTCTGCT 1092
Qy 4449 TAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4508
Db 1093 T-AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1151
Qy 4509 CAACATGTTTAAAGATGTGAGAGCTCTGCTTCCATGCTGTAAAAAACCACCTCA 4568
Db 1152 CAACATGTTTAAAGATGTGAGAGCTCTGCTTCCATGCTGTAAAAAACCACCTCA 1211
Qy 4569 AGGCGAGGTGAGTGGCTCATGCGCTATATATCCAGCATTTTGGGAGCGGAGGAGGAG 4628
Db 1212 AGGCGAGGTGAGTGGCTCATGCGCTATATATCCAGCATTTTGGGAGCGGAGGAGGAG 1271
Qy 4629 ATCACTGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4688
Db 1272 ATCACTGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331
Qy 4689 TACTAAAAATACAAATTTAGATGAGCATGTTGTGTCATGCTGTATCCACCTACTTGG 4748
Db 1332 TACTAAAAATACAAATTTAGATGAGCATGTTGTGTCATGCTGTATCCACCTACTTGG 1391
Qy 4749 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4808
Db 1392 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
Qy 4809 GTGCCATTGCACT-CCAGCCTGAGCA-ATGAGCGAAATCCATCTCAAAAAAAGCAAC 4866
Db 1452 GTGCCATTGCACTNCCAGCCTGAGCANAATGAGCGAAATCCATCTCAAAAAAAGCAAC 1511
Qy 4867 AAAAAACCACTCTTACTCCAGGAGGAGTGGGTCA- GAGCTGGGCGCATCATGTCAG 4925
Db 1512 AAAAAACCACTCTTACTCCAGGAGGAGTGGGTCA- GAGCTGGGCGCATCATGTCAG 1571
Qy 4926 GTGCTGAGGCGCAGAGGTAAGGCGGAGCTG- -CAGGACCGGCGGACAGATTAACGTGCT 4983
Db 1572 GTGCTGAGGCGCAGAGGTAAGGCGGAGCTG- -CAGGACCGGCGGACAGATTAACGTGCT 1631
Qy 4984 GAGATCAGTGTGTAGATCAGACGTCCTCTGTCATTTGTGATCACCAGGAGGCGCCCAAGC 5043
Db 1632 GAGATCAGTGTGTAGATCAGACGTCCTCTGTCATTTGTGATCACCAGGAGGCGCCCAAGC 1691
Qy 5044 ACCGAGATGGCCCCCATCCAGTCAACACATCCATCTTCATCCAGAGATGTCTGTTCTT 5103
Db 1692 ACCGAGATGGCCCCCATCCAGTCAACACATCCATCTTCATCCAGAGATGTCTGTTCTT 1751
Qy 5104 GGCACGCTGGGGTAAATTAGACAGAGGTGACAGTCTGGGTGGGTCATTCAGACTGC 5163
Db 1752 GGCACGCTGGGGTAAATTAGACAGAGGTGACAGTCTGGGTGGGTCATTCAGACTGC 1811
Qy 5164 CCCAGGCA- -GGCCTGTGGCCTGT- -AGAAAAAGTTAGGCGCTAGGCGGCGCAGCGT 5217
|||||

Db 1812 CCCAGGNNCAGGNCCTNTGTGGCTGNTANGAAAAAGTTACAGGCTAGGCGGACGCT 1871
Qy 5218 GGC-TCAGGCTGTATATCCAGCACTTTTGGGAGGCGGAGGCGGTGATCAAGAGTCA 5276
Db 1872 GGCNTCAGGCTGTATATCCAGCACTTTTGGGAGGCGGAGGCGGTGATCAAGAGTCA 1931
Qy 5277 GAGATCGTACCATCTGCTTAACAGCGGAAACCCGCTCTACTTAAAAATACAAAAA 5336
Db 1932 GAGATCGTACCATCTGCTTAACAGCGGAAACCCGCTCTACTTAAAAATACAAAAA 1991
Qy 5337 TTGGCGGCGCATGTGGCGGCGGACCTGTATTCAGTACTGCGGAGGCTGAGGAGAG 5396
Db 1992 TTGGCGGCGCATGTGGCGGCGGACCTGTATTCAGTACTGCGGAGGCTGAGGAGAG 2051
Qy 5397 AATGCGGTGAAACCCAGAGGCGAGATTTTGCATGAGCCGGAATTCGCGCACCTGCA 5456
Db 2052 AATGCGGTGAAACCCAGAGGCGAGATTTTGCATGAGCCGGAATTCGCGCACCTGCA 2111
Qy 5457 GCCTGGGCGACAGACAGACATCCATCTGGA- -AGAAAAAGAAAAAGTTACAGTCTGAG 5515
Db 2112 GCCTGGGCGACAGACAGACATCCATCTGGA- -AGAAAAAGAAAAAGTTACAGTCTGAG 2171
Qy 5516 CCAGAGGCCAGGCTGTATTTCTGTCATTCATGACCTTGGGCAAGGCACTTCTTC 5575
Db 2172 CCAGAGGCCAGGCTGTATTTCTGTCATTCATGACCTTGGGCAAGGCACTTCTTC 2231
Qy 5576 CTGGCCGAGTTTACGCGGTTGGAATGACCTCAAGGTCCTTCCAGCATTAACGTCAT 5635
Db 2232 CTGGCCGAGTTTACGCGGTTGGAATGACCTCAAGGTCCTTCCAGCATTAACGTCAT 2291
Qy 5636 GGTCT- AAGATGAAAGATGGGAGATTTCCCTCTGTCACCCGAGCCGCTGACAT 5694
Db 2292 GGTCT- AAGATGAAAGATGGGAGATTTCCCTCTGTCACCCGAGCCGCTGACAT 2351
Qy 5695 CAAGGTGATGACAGGAGGAGTCACTCCCAATCCGCGAGTTCCAAAGCCCTTGG 5752
Db 2352 CAAGGTGATGACAGGAGGAGTCACTCCCAATCCGCGAGTTCCAAAGCCCTTGG 2409

RESULT 14
US-10-217-651-592
; Sequence 592, Application US/10217651
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P2121CIN
; CURRENT APPLICATION NUMBER: US/10/217,651
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/760,491
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647

PRIOR FILING DATE: 2000-07-07	PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07	PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08	PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27	PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21	PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21	PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30	PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26	PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20	PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20	PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01	PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08	PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08	PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25	PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05	PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08	PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05	PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-13	PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13	PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13	PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20	PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08	PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08	PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22	PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22	PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27	PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06	PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30	PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14	PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08	PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08	PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08	PRIOR APPLICATION NUMBER: 60/233,064

1	PRIOR FILING DATE: 2000-09-14
2	PRIOR APPLICATION NUMBER: 60/233,063
3	PRIOR FILING DATE: 2000-09-14
4	PRIOR APPLICATION NUMBER: 60/233,397
5	PRIOR FILING DATE: 2000-09-14
6	PRIOR APPLICATION NUMBER: 60/233,399
7	PRIOR FILING DATE: 2000-09-14
8	PRIOR APPLICATION NUMBER: 60/233,401
9	PRIOR FILING DATE: 2000-09-14
10	PRIOR APPLICATION NUMBER: 60/241,808
11	PRIOR FILING DATE: 2000-10-20
12	PRIOR APPLICATION NUMBER: 60/241,826
13	PRIOR FILING DATE: 2000-10-20
14	PRIOR APPLICATION NUMBER: 60/241,786
15	PRIOR FILING DATE: 2000-10-20
16	PRIOR APPLICATION NUMBER: 60/241,221
17	PRIOR FILING DATE: 2000-10-20
18	PRIOR APPLICATION NUMBER: 60/246,475
19	PRIOR FILING DATE: 2000-11-08
20	PRIOR APPLICATION NUMBER: 60/231,243
21	PRIOR FILING DATE: 2000-09-08
22	PRIOR APPLICATION NUMBER: 60/233,065

Query Match	35.7%	Score 2225.8;	DB 48;	Length 2409;
Best Local Similarity	98.7%;	Pred. No. 1.6e-257;		
Matches 2367; Conservative	0;	Mismatches 14;	Indels 17;	Gaps 12;

QY	3371	CCCTGTCAAAAACAGACACACCTCGGCTCTACCAATAGACAGATG-TGGGTGGGGCCACATA	34299
Db	13	CCNCTGTAAAAACACACACTCGGCTCTACCAATAGACAGATGATGATGGGTGGGGCCAGATA	72
QY	3430	AGAGAAATAAAAAGCAGGCTGCCCGAGCCAGCAGTGGCAACGGCGCAGGTCCCTTATCCACA	3489
Db	73	AGAGAAATAAAAAGCAGGCTGCCCGAGCCAGCAGTGGCAACGGCGCAGGTCCCTTATCCACA	132
QY	3490	ATATGCGAGCTTGTGTCCTTTGCTTGCTTGGCATTAATCTTGCTACTGCTGCTTTT-TGG	3548
Db	133	ATATGCGAGCTTGTGTCCTTTGCTTGCTTGGCATTAATCTTGCTACTGCTGCTTTTMMG-192	
QY	3549	GTCACACCTGCTTTTATGAGCTGTAAACATCTACACAGAAAGTCTGCACTTCACTCTGA	3608
Db	193	GTCACACCTGCTTTTATGAGCTGTAAACATCTACACAGAAAGTCTGCACTTCACTCTCGA	252
QY	3609	AGCCACTTAAGACCCAGAGCCACCAGGGAGGAATGAACAATCCGCGCGGCTGTCCTTAAG	3668
Db	253	AGCCACTTAAGACCCAGAGCCACCAGGGAGGAATGAACAATCTCCGGCGCGCTGCTTAAAG	312
QY	3669	AGCTATTAACACTCCGCGGAAGGTCCTGACGCTTCACTCTCAGCCAGCAGAACCCAGCAAC	3728
Db	313	AGCTATTAACACTCCGCGGAAGGTCCTGACGCTTCACTCTCAGCCAGCAGAACCCAGCAAC	372
QY	3729	CCACCAAAAGAGAAAACTGCGAACAATCTGAAACATTCAGAGAACAACTCCAGATGC	3788
Db	373	CCACCAAAAGAGAAAACTGCGAACAATCTGAAACATTCAGAGAACAACTCCAGATATGC	432
QY	3789	ACCACCTTAAAGAGTGTAAACATCTACGCGAGGTCGCGCGCTTCCTTCTTGAAGTCAGT	3848
Db	433	ACCACCTTAAAGAGTGTAAACATCTACGCGAGGTCGCGCGCTTCTTCTTGAAGTCAGT	492
QY	3849	GAGACCAAGCACTCACCAAGTTTGGACACAAGCCAGAGATTTAGATTCAGCTCGGGCAA	3908
Db	493	GAGACCAAGCACTCACCAAGTTTGGACACAAGCCAGAGATTTAGATTCAGCTCGGGCAA	552
QY	3909	CATGATTAATTAATGCCCTCTCTGCAAAAAAATTAACAAAAATTTGGGGAGCATGCT	3968
Db	553	CATGATTAATTAATGCCCTCTCTGCAAAAAAATTAACAAAAATTTGGGGAGCATGCT	612
QY	3969	GGTCCGTGCGCTGTGGTCCCAAGTACGCGGGAGAGCTTAAGTGGAGAGATCGCTTAAGCCTG	4028
Db	613	GGTCCGTGCGCTGTGGTCCCAAGTACGCGGGAGAGCTTAAGTGGAGAGATCGCTTAAGCCTG	672
QY	4029	GGAGGTGAAGCTGCACTGAGCTGTGTGATTGACCAACGCCCTTAGGCTGGGGAGACAG	4088

Db	673	GGAGGTGAAGACTGCAGTGAAGCTGTGATTGTACCAAGCCCTCTAAGCTGGGGGACAGAC	732
OY	4089	TGAGACCCCTGTTTCCCTCCGCAAAAAATTGACAAAAGTGTATTAAGAGTGGCTGTATA	4148
Db	733	TGAGACCCCTGTTTCCCTCCGCAAAAAATTGACAAAAGTGTATTAAGAGTGGCTGTATA	792
OY	4149	TGGCTAGGCGCAGTGGCTCATGCTGTATATCCACAGACTTTGGGAGCCGAGGGGGGG	4208
Db	793	TGGCTAGGCGCAGTGGCTCATGCTGTATATCCACAGACTTTGGGAGCCGAGGGGGGG	852
OY	4209	GTCACCTTAAGGTCAGGAGTGTGAGACCGAGCTGGGCGAACAATGGAGAAAGCCATCTCTC	4268
Db	853	GTCACCTTAAGGTCAGGAGTGTGAGACCGAGCTGGGCGAACAATGGAGAAAGCCATCTCTC	912
OY	4269	TAAAAATACAAAATTAGCCGGCTGTGGGGGCGAGTGGAGACATGACCCTGTATCCAGCT	4328
Db	913	TAAAAATACAAAATTAGCCGGCTGTGGGGGCGAGTGGAGACATGACCCTGTATCCAGCT	972
OY	4339	ACTCAGAGGCTGAGCGAGAGAACTCACTTGAACCCAGAGGGCGGGTTCAGTGAACC	4388
Db	973	ACTCAGAGGCTGAGCGAGAGAACTCACTTGAACCCAGAGGGCGGGTTCAGTGAAGCC	1032
OY	4389	GAGATCGTGCATTTGCCACTCCACCCACTCCAGCCCTGGGCAACAAGGCCAAACTCTGCT	4448
Db	1033	GAGATCGTGCATTTGCCACTCCACCCACTCCAGCCCTGGGCAACAAGGCCAAACTCTGCT	1092
OY	4449	TAAAAAATAAAAAAAAAAGTGCCTGTACATTAAGAGTGTGCATATGCATATGTTGGCCAGG	4508
Db	1093	T-AAAAAAAAAAAAAAAAAGTGCCTGTACATTAAGAGTGTGCATATGCATATGTTGGCCAGG	1151
OY	4509	CAACATGTTTAAGAAATGTGAGATCTGCTCCCTTCATGGTCTCTGTTAAAAAACCCACTCA	4568
Db	1152	CAACATGTTTAAGAAATGTGAGATCTGCTCCCTTCATGGTCTCTGTTAAAAAACCCACTCA	1211
OY	4569	AGGCGAGGTGAGTGGCTCATGCTGTATATCCACAGACTTTGGGAGGCCGAGGGGGTGG	4628
Db	1212	AGGCGAGGTGAGTGGCTCATGCTGTATATCCACAGACTTTGGGAGGCCGAGGGGGTGG	1271
OY	4629	ATCACCTTAGGTCAGGAGTTGTGAGACCGAGCCGTGACACACATATGTTAAATCCACCTC	4688
Db	1272	ATCACCTTAGGTCAGGAGTTGTGAGACCGAGCCGTGACACACATATGTTAAATCCACCTC	1331
OY	4689	TACTAAAAATACAAAATTAGATGAGCAGATGTGTGATGCTGTATATCCCACTACTTGG	4748
Db	1332	TACTAAAAATACAAAATTAGATGAGCAGATGTGTGATGCTGTATATCCCACTACTTGG	1391
OY	4749	GAGCGTGAAGCAGGAAAAATCACTGTAACACAGGAGGGGAGGTTGTAGTAGCCGAGATC	4808
Db	1392	GAGCGTGAAGCAGGAAAAATCACTGTAACACAGGAGGGGAGGTTGTAGTAGCCGAGATC	1451
OY	4809	GTCGCATTTGCACT--CCAGCCGTGACA--ATGAGCCAAACTCCATCTCAAAAAACAACAC	4866
Db	1452	GTCGCATTTGCACTCCACCTGTGACACANATGAGCAAACTCCATCTCAAAAAACAACAC	1511
OY	4867	AAAAAACCCACCTCTACTCCAGGAGGAGTGGGTACA--GAGTGGGGCCATCACTGCAAG	4925
Db	1512	AAAAAACCCACCTCTACTCCAGGAGGAGTGGGTACANAGAGTGGGCCATCACTGCAAG	1571
OY	4926	GTGCTGAAGCCACAGAGCTTAAGCGGAGCTG--CAGAGCCGGGACAGATTAACAGTGTGT	4983
Db	1572	GTGCTGAAGCCACAGAGCTTAAGCGGAGCTGNNCACAGACCGCGGACAGATTAACAGTGTGT	1631
OY	4984	GAGATTCAGTGTGTAGATCAGAGCTCCCTGCAATTTGTGACACACAGGGGGCCCCAAGC	5043
Db	1632	GAGATTCAGTGTGTAGATCAGAGCTCCCTGCAATTTGTGACACACAGGGGGCCCCAAGC	1691
OY	5044	ACCAAGAGTGGCCCATATCCAGTCAACACATCCACTTGTCAATCCAGAGATGTCTGTTCTT	5103
Db	1692	ACCAAGAGTGGCCCATATCCAGTCAACACATCCACTTGTCAATCCAGAGATGTCTGTTCTT	1751
OY	5104	GGCAGCGCTGGGTAATTAAGACAGAAAGTGACAGTCTTGGGTGTGTCAGTCAAGACTGC	5163
Db	1752	GGCAGCGCTGGGTAATTAAGACAGAAAGTGACAGTCTTGGGTGTGTCAGTCAAGACTGC	1811

```

QY 5164 CCGAGGCA-----GGCCTTGTGGCCTGT-AGAAAACTTCAGGCTTAGGCCGGGACGGT 5217
    |||||      |      |      |      |      |      |      |      |      |      |
Db 1812 CCGAGGNNCAGGNNCTNTGTGGCCTGTNTANGAAAACTTTCAGGCTTAGGCCGGGACGGT 1871
QY 5218 GGC-CGACGGCTGTAAATCCAGACACTTTGGGAGGGCGAGGGGGGGTGGATCAGAGTCAAG 5276
    |||      |      |      |      |      |      |      |      |      |      |
Db 1872 GGCNTCAGCCCTGTATATCCAGCACACTTTGGGAGGGCGAGGGGGGATCAGAGTCAAG 1931
QY 5277 GAGATCGTGACCATCTCTGGCTTAACACGGGTAAACCCCGCTCTTACTTAAATAATCAAAAAA 5336
    |||||      |      |      |      |      |      |      |      |      |      |
Db 1932 GAGATCGTGACCATCTCTGGCTTAACACGGGTAAACCCCGCTCTTACTTAAATAATCAAAAAA 1991
QY 5337 TTGGCCGGGGCATGTGTGGCGGGCACCTGTATGTTCCAGCTACTTCGGGAGGCTGAGGACGAG 5396
    |||||      |      |      |      |      |      |      |      |      |      |
Db 1992 TTGGCCGGGGCATGTGTGGCGGGCACCTGTATGTTCCAGCTACTTCGGGAGGCTGAGGACGAG 2051
QY 5397 AATGGCGTGAACCCGAGAGGACGAGATTTGTGCANTGAGCCGAGATCGCGCACATGCTCCA 5456
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2052 AATGGCGTGAACCCGAGAGGACGAGATTTGTGCANTGAGCCGAGATCGCGCACATGCTCCA 2111
QY 5457 GCCTTGGCGACAGAGACACTCATCTGGAAA-AGAAAAAAGAAAGCTTCAGGCTGAG 5515
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2112 GCCTTGGGGACAGACAGACACTCATCTGGAAAAGAAAGAAAGAAAGCTTCAGGCTGAG 2171
QY 5516 CCAGAGGCCACAGGCTGTATTTCTGTCACTTACCATGACCTTGGGGAAGGCACTTCTTCC 5575
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2172 CCAGAGGCCACAGGCTGTATTTCTGTCACTTACCATGACCTTGGGGAAGGCACTTCTTCC 2231
QY 5576 CTGGGCCAGTTTCAGGGGTTGGAAATCGACTCCAAAGTCCCTTCAGACATTAACGCTGCAT 5635
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2232 CTGGGCCAGTTTCAGGGGTTGGAAATCGACTCCAAAGTCCCTTCAGACATTAACGCTGCAT 2291
QY 5636 GGTTCCT-AAGATGAGAAGATGGGGCAGTTTCCCTCTCTCACCCAGCCCGTGTCACTT 5694
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2292 GGTTCCTAAGATGAGAAGATGGGGCAGTTTCCCTCTCTCACCCAGCCCGTGTCACTT 2351
QY 5695 CAAAGTGAAATGACACAGGGAAAGTCAAGTGTCCCAATCCGCAATTCGCAAGCCCTTGGG 5752
    |||||      |      |      |      |      |      |      |      |      |      |
Db 2352 CAAAGTGAAATGACACAGGGAAATCAAGTGTCCCAATCCGCAATTCGCAAGCCCTTGGG 2409
    |||||      |      |      |      |      |      |      |      |      |      |

RESULT 15
US-60-466-412-86428/c
; Sequence 86428, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO011466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86428
; LENGTH: 46397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(46397)
; OTHER INFORMATION: n = A,T,C or G
US-60-466-412-86428

```

Query Match	31.4%;	Score 1959.4;	DB 101;	Length 46997;
Best Local Similarity	99.8%;	Pred. No. 6.5e-226;		
Matches 2004: Conservative	0;	Mismatches 1;	Indels 4;	Gaps 4;

QY 4226 GFTTGAGACCAAGCTTGGCCCAACATGGAGAAAGCCCACTCTCTTCTTAAATAACAAATTAG 4285

Db 46997 GTGTGAGACCAAGCTTGGCCCAACATGGAGAAAGCCCACTCTCTTCTTAAATAACAAATTAG 46938

QY	4286	CCGGCTGTGGGGGAGAGTGGTGGACATGCCCTGTAAATCCACAGCTACTACAGAGGCTGAGGC	4345
Db	46937	CCGGCTGTGGGGGAGAGTGGTGGACATGCCCTGTAAATCCACAGCTACTACAGAGGCTGAGGC	4687
QY	4346	AGAGAAATACATTGAAACCCAGAGGCGGGTTGCAGTACGAGACCGAGATCGGCAATTGCA	4405
Db	46877	AGGAAATACATTGAAACCCAGAGGCGGGTTGCAGTACGAGACCGAGATCGGCAATTGCA	4681
QY	4406	CTCCACCCACTCCAGCCCTGGGGCAACAGAGCCAAACTCTGTCTTAAAAA	4465
Db	46817	CTCCACCCACTCCAGCCCTGGGGCAACAGAGCCAAACTCTGTCTTAAAAA	4675
QY	4466	AGTGCCTTACATATTAAGAGTGTGCATATGCTATGTTCCAGGCAACATGTTTAAAGATG	4525
Db	46758	AGTGCCTTACATATTAAGAGTGTGCATATGCTATGTTCCAGGCAACATGTTTAAAGATG	4669
QY	4526	TGGAGCTCTGCCTTCCATGGTCCGTAAAAACCCACCCTCAAGGGCAGGTGCGAGTGGC	4585
Db	46688	TGGAGCTCTGCCTTCCATGGTCCGTAAAAACCCACCCTCAAGGGCAGGTGCGAGTGGC	4663
QY	4586	TCATGCTCATATATCCACAGCACTTTTGGAGGCGGAGGGGGTGTATCATCTGAGTCAAGA	4645
Db	46638	TCATGCTCATATATCCACAGCACTTTTGGAGGCGGAGGGGGTGTATCATCTGAGTCAAGA	4657
QY	4646	GTTGAGACACACGCTGACCACCAACATGCTGAATCCCACTCTACTAAAAATACAAAT	4705
Db	46578	GTTGAGACACACGCTGACCACCAACATGCTGAATCCCACTCTACTAAAAATACAAAT	4651
QY	4706	TAGATGACATAGTGTGTCATGCTCTGTAAATCCCACTACTTGGAGGCTGAGGAGAAA	4765
Db	46518	TAGATGACATAGTGTGTCATGCTCTGTAAATCCCACTACTTGGAGGCTGAGGAGAAA	4645
QY	4766	ATCACTAGAAACAGAGGAGGCGAGAGTGTGATGAGCGAGATCGTGCATTTGCACTCAG	4825
Db	46648	ATCACTAGAAACAGAGGAGGCGAGAGTGTGATGAGCGAGATCGTGCATTTGCACTCAG	4639
QY	4826	CCTGAGCAATGAGCGAAATCTCATCTCAAAAAAACAACAACAAACCCACTCTCACTC	4885
Db	46398	CCTGAGCAATGAGCGAAATCTCATCTCAAAAAAACAACAACAAACCCACTCTCACTC	4633
QY	4886	CCAGGAGCTGGGTACAGAGCTGGGGCAACATCAAGTGAAGGTGCTGAGCCACAGAGTAA	4945
Db	46338	CCAGGAGCTGGGTACAGAGCTGGGGCAACATCAAGTGAAGGTGCTGAGCCACAGAGTAA	4627
QY	4946	GGCGAGCTGCAGAGACCGCGAGCAACAGATAAACAGTGTGTGAGATCAAGTGTGAGATAGA	5005
Db	46278	GGCGAGCTGCAGAGACCGCGAGCAACAGATAAACAGTGTGTGAGATCAAGTGTGAGATAGA	4621
QY	5006	CGTCCCTGCAATTTGGTACACACAGGGGGCCCCAGACCAACAGATGGCCCCATCCAGT	5065
Db	46218	CGTCCCTGCAATTTGGTACACACAGGGGGCCCCAGACCAACAGATGGCCCCATCCAGT	4615
QY	5066	CACCAACATCCACTCTCATCCAGAGATGTCTGTTCTTGGACAGCTGGGGTAAATTGGA	5125
Db	46158	CACCAACATCCACTCTCATCCAGAGATGTCTGTTCTTGGACAGCTGGGGTAAATTGGA	4609
QY	5126	CAGAAGGTGAAGTCTTGGGTGTGATCAGATCAGACTCCCCAGGCAAGCCCTTGGGCGTG	5185
Db	46098	CAGAAGGTGAAGTCTTGGGTGTGATCAGATCAGACTCCCCAGGCAAGCCCTTGGGCGTG	4603
QY	5186	TAGAAAACGTTTCAAGGCTTAGGCGGGGACAGGTTGGCTCAAGCCTGTAAATCCACGACTTGG	5245
Db	46038	TAGAAAACGTTTCAAGGCTTAGGCGGGGACAGGTTGGCTCAAGCCTGTAAATCCACGACTTGG	4597
QY	5246	GGAGGCGGAGGCGGGTGGATACACAGATCAGAGATGTGACCATCTGGCTAACACGGT	5305
Db	45978	GGAGGCGGAGGCGGGTGGATACACAGATCAGAGATGTGACCATCTGGCTAACACGGT	4591
QY	5306	GAAACCCCGTCTACTATTAATAATCAAAAAATTGGCGGGGCAATGTGGCGGGGACCGCTGA	5365
Db	45918	GAAACCCCGTCTACTATTAATAATCAAAAAATTGGCGGGGCAATGTGGCGGGGACCGCTGA	4585
QY	5366	GTTCACGACTACTCCGGAGGCTGAGGCGAGGAAATGGCGTGAACCCGAGAGCAAGTTGG	5425

Db 45858 GTTCAGCTACTCGGAGGCTGAGGACAGAGAAATGGCTGAAACCGAGAGGAGATTG 45799
QY 5426 CAGTGAGCCGAGATCGGCCCTGCATCCAGCCTGGGGGACAGAGCAAGACTCCATCTG 5485
Db 45798 CAGTGAGCCGAGATCGGCCCTGCATCCAGCCTGGGGGACAGAGCAAGACTCCATCTG 45739
QY 5486 GAAAGAGAAAGAAAGCTTACGCTGAGCCAGAGGCCAGGCTGTAATCTGTCACTT 5545
Db 45738 GAAAGAGAAAGAAAGCTTACGCTGAGCCAGAGGCCAGGCTGTAATCTGTCACTT 45679
QY 5546 ACCATGACCTTGGGCAAGGACCTTCCTCCCTGGGCCAGTTCACGGGGTTGGAATCGACT 5605
Db 45678 ACCATGACCTTGGGCAAGGACCTTCCTCCCTGGGCCAGTTCACGGGGTTGGAATCGACT 45619
QY 5606 CCAAGGTCCTTCCAGCATTAACCTGATGCTCTAGATGAGAGATGGGGCATTTTC 5665
Db 45618 CCAAGGTCCTTCCAGCATTAACCTGATGCTCTAGATGAGAGATGGGGCATTTTC 45559
QY 5666 CCGTCTCTACCCCGAGCCCGTGTCCACTTCAAGTGATGACACAGGAAGTCACTGTCC 5725
Db 45558 CCGTCTCTACCCCGAGCCCGTGTCCACTTCAAGTGATGACACAGGAAGTCACTGTCC 45499
QY 5726 CAATCCGCACTTCCAAAGCCCTTGGGGACCTACTGTCAAGGTCGTGCACGAGAGGTG 5785
Db 45498 CAATCCGCACTTCCAAAGCCCTTGGGGACCTACTGTCAAGGTCGTGCACGAGAGGTG 45439
QY 5786 AAGGTCAGGTGAGCCATCGGCGAAGGGTCTGCTCATTCGGGACAGACATCCGGTT 5845
Db 45438 AAGGTCAGGTGAGCCATCGGCGAAGGGTCTGCTCATTCGGGACAGACATCCGGTT 45379
QY 5846 TCCTCTGCTTACCGGATTCAGGGGCTTAGCCGAATGATCATGGGGGCGGGGG 5905
Db 45378 TCCTCTGCTTACCGGATTCAGGGGCTTAGCCGAATGATCATGGGGGCGGGGG 45319
QY 5906 GTTCTGGGGAGTTCACGCTAATCAACTTGGGACAGAGACCCCTGGAACCTTGCATGG 5965
Db 45318 GTTCTGGGGAGTTCACGCTAATCAACTTGGGACAGAGACCCCTGGAACCTTGCATGG 45259
QY 5966 TGCTATCCAGTGCGGGGTGGGACAGCAGCAAGACCAATGCTTATCTCAGGTAG 6025
Db 45258 TGCTATCCAGTGCGGGGTGGGACAGCAGCAAGACCAATGCTTATCTCAGGTAG 45199
QY 6026 GGGCTCAGAGGCTCCAGACAGGACCTCCGAGAGTTTGGGGGTAGGATGGAGC 6085
Db 45198 GGGCTCAGAGGCTCCAGACAGGACCTCCGAGAGTTTGGGGGTAGGATGGAGC 45139
QY 6086 AACGAGCTTCTTTTCTCTTAGAATTTGGGGCTTGGGGACAGGCTTGAATTC 6145
Db 45138 AACCA-GCTTCTTTTCTCTTAGAATTTGGGGCTTGGGGACAGGCTTGAATTC 45080
QY 6146 CCAAGAGAGAGGGGCAAGGACACT-CCCCCAAGTGTCCAGAGCGAGAGGAGAGC 6204
Db 45079 CCAAGAGAGAGGGGCAAGGACACTGCCCCGCAAGTGTCCAGAGC-AGAGAGGAGAGC 45021
QY 6205 CCGGACTCAGCTGCACCTTCCCAAGGC 6233
Db 45020 CCGGACTCAGCTGCACCTTCCCAAGGC 44992

Search completed: August 17, 2003, 01:29:01
Job time : 11330 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	5532.6	88.7	14381	7	US-60-487-610-19745	Sequence 19715, A
2	942	15.1	190672	7	US-60-487-610-19215	Sequence 19215, A
3	860	13.8	1748339	5	US-09-947-914-48	Sequence 48, Appl
4	767	12.3	30105	7	US-60-487-610-19434	Sequence 19434, A
5	767	12.3	30105	7	US-60-485-450-11978	Sequence 11978, A
6	752	12.1	108815	7	US-60-487-610-19424	Sequence 19424, A
7	752	12.1	108815	7	US-60-485-450-11975	Sequence 11975, A
8	751.2	12.0	1383163	5	US-09-947-914-41	Sequence 41, Appl
9	730.6	11.8	121924	6	US-10-450-826-50	Sequence 50, Appl
10	720.2	11.6	29405	7	US-60-487-610-19466	Sequence 19466, A
11	710.4	11.4	13040	7	US-60-485-450-11933	Sequence 11933, A
12	710.4	11.4	48853	7	US-60-485-450-11934	Sequence 11934, A
13	693	11.1	40207	7	US-60-487-610-20017	Sequence 20017, A
14	677.2	10.9	70556	7	US-60-487-610-19239	Sequence 19239, A
15	668	10.7	185371	6	US-10-470-050-852	Sequence 852, App
16	636.8	10.2	71087	7	US-60-487-610-20097	Sequence 20097, A
17	633.4	10.2	60717	7	US-60-487-610-20086	Sequence 20086, A
18	633.4	10.2	60717	7	US-60-485-450-12341	Sequence 12341, A
19	631	10.1	113331	7	US-60-485-450-12302	Sequence 12302, A
20	626.2	10.0	37027	7	US-60-485-450-12084	Sequence 12084, A
21	607.6	9.7	1383163	5	US-09-947-914-41	Sequence 41, Appl
22	602.6	9.7	352800	7	US-60-487-610-19447	Sequence 19447, A
23	598.2	9.6	544420	7	US-60-487-610-19233	Sequence 19233, A
24	584.2	9.4	882900	6	US-10-292-798-13393	Sequence 1333, Appl
25	578.2	9.3	8059021	5	US-09-947-914-53	Sequence 93, Appl
26	574.8	9.2	21329	7	US-60-487-610-19363	Sequence 19363, A

27	573.8	9.2	21331	7	US-60-487-610-19358	Sequence 19358, A
28	573.8	9.2	1701331	5	US-60-487-610-19754	Sequence 46, A
29	564.4	9.1	41578	7	US-09-947-914-46	Sequence 19754, A
30	559	9.0	8059021	5	US-09-947-914-45	Sequence 53, A
31	557.2	8.9	1946141	5	US-09-947-914-44	Sequence 42, A
32	554	8.8	349443	5	US-09-947-914-47	Sequence 47, A
33	547.8	8.8	820835	5	US-60-487-610-19793	Sequence 47, A
34	533.6	8.6	169739	6	US-10-450-826-63	Sequence 19799, A
35	530	8.5	12253	7	US-60-485-450-12194	Sequence 12194, A
36	520.4	8.3	16868	7	US-60-485-450-11866	Sequence 11866, A
37	515	8.3	342574	7	US-60-485-450-11974	Sequence 11974, A
38	502.8	8.1	32015	7	US-60-487-610-19671	Sequence 19671, A
39	499.6	8.1	33023	7	US-60-485-450-12181	Sequence 12181, A
40	492.8	8.0	74079	7	US-60-487-610-19261	Sequence 19261, A
41	498.4	8.0	377157	7	US-60-487-610-19429	Sequence 19429, A
42	490	7.9	37637	7	US-60-485-450-11859	Sequence 11859, A
43	487.2	7.8	127361	7	US-60-485-450-11857	Sequence 11857, A
44	469.4	7.5	47689	7	US-60-485-450-12330	Sequence 12330, A
45	465	7.5	94564	7	US-60-487-610-19683	Sequence 19683, A

ALIGNMENTS

```

RESULT 1
US-60-487-610-19745
: Sequence 19745, Application US/60487610
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: HOANG, Hongjin
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-
: TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001469
: CURRENT APPLICATION NUMBER: US/60/487,610
: CURRENT FILING DATE: 2003-07-17
: NUMBER OF SEQ ID NOS: 97101
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 19745
: LENGTH: 14381
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-487-610-19745

```

Query Match	88.78;	Score 5532.6;	DB 7;	Length 14381;
-------------	--------	---------------	-------	---------------

```

Best Local Similarity: 55.58; Freq: NO: 0;
Matches 5655; Conservative 0; Mismatches 14; Indels 14; Gaps 10;

```

QY	558	AAAAAAGAGCTGCTTTAATGTCGTCGCGCCATACATACGTAAGTATATAGTTTGCA	617
Db	1	AAAAAAGAGCTGTTTAAATGTCCTCCGGTACATCA	57
		----	-----TACTAATGATAAAGTTTCCAA
QY	618	ACTCAAGAAGCCAGATAGTCAATTTTTTATGGCTTGCGCGCGTATAGTCTGTGCACAT	677
Db	58	ACTCAAGAAGTCCAGATAGTCAATTTTTTATGGCTTGCGCGCGTATAGTCTGTGCACAT	117
QY	678	CACCTGCCCCGTGCTTTTACGACAAAGACAGCTATTAACATACATCATGAATTTTTT	737
Db	118	CACCTGCCCCGTGCTTTTACGACAAAGACAGCTATTAACATACATCATGAATTTTTT	177
QY	738	ATAGACATCGAGATTTGAATTCATATGATTTTTTACATTTTATATAAATATCTTTTAAA	797
Db	178	ATAGACATCGAGATTTGAATTCATATGATTTTTTACATTTTATATAAATATCTTTTAAA	237
QY	798	AATTTTCCCCCTAACATTTTAAAAGTATAAAGCGCGCA	858
		-----GCGCGCATCGTCACGCC	918
Db	238	AATTTTCCCCCTAACATTTTAAAAGTATAAAGCGCGCGCGCCATGGCTTCACGCC	297
QY	855	TGTAAATTCAGACCTTTGGGAGGCGAGGTGGCGAGATCACTTGGATCAACACTTTCGAG	914
Db	298	TGTAAATTCAGACCTTTGGGAGGCGAGGTGGCGAGATCACTTGGATCAACACTTTCGAG	357
QY	915	ACGAGCGTGGCCAAACATATACAAAAACCCCATTTCTACTAAAAATATAAAAAATTTAGCTGGGC	974

|||||
Db ACCAGCGTGGCAACATAGAAAAACCCATTCTACTAAAAATAAAAATTTAGCTGGC 417
QY ATATGGTGCACACCTGTGATCCAGCTACTTGGAGCGTGAGCGACGAGAATCGTTGA 1034
Db ATATGGTGCACACCTGTGATCCAGCTACTTGGAGCGTGAGCGACGAGAATCGTTGA 477
QY ACCTGGGAACGGAGGTTGAGTGAAGCAATCATCTCCACTCCAGCCCTGGGTGA 1094
Db ACCTGGGAACGGAGGTTGAGTGAAGCAATCATCTCCACTCCAGCCCTGGGTGA 537
QY CAGAGTGAGACTTCTCTCAACGAAAAAAAGTGTAAAAAGCCATTCTTAATTCAGTGT 1154
Db CAGAGTGAGACTTCTCTCAACGAAAAAAAGTGTAAAAAGCCATTCTTAATTCAGTGT 597
QY ACATCAGTGTACATCTCAGGCTCGCTACTCTCTGTGAGGCATACCTGAGAAGTGA 1214
Db ACATCAGTGTACATCTCAGGCTCGCTACTCTCTGTGAGGCATACCTGAGAAGTGA 657
QY GTTGTGTTGGTCAAGGACATACATTTTCCATTTAATCTAGACATACCAAGTTGCCATC 1274
Db GTTGTGTTGGTCAAGGACATACATTTTCCATTTAATCTAGACATACCAAGTTGCCATC 717
QY CAAGGAGG-TTTTTTTTTTACATCTACACTCCCCCAGCAGCAAAATGAGAGTTACTCA 1333
Db CAAGGAGGTTTTTTTTTACATCTACACTCCCCCAGCAGCAAAATGAGAGTTACTCA 777
QY GATCCTTTACAAAGATGCTCTAAGCCAGTACAGATGAAAACAGGAAGTGGAGGGAA 1393
Db GATCCTTTACAAAGATGCTCTAAGCCAGTACAGATGAAAACAGGAAGTGGAGGGAA 837
QY GGTGGCAGGCCCTTCTAACCATGAGAAATACCTGTGAGAGCCTTCTGTGATGCTGGAAG 1453
Db GGTGGCAGGCCCTTCTAACCATGAGAAATACCTGTGAGAGCCTTCTGTGATGCTGGAAG 897
QY ATGAATAGGAGGCGTCTGAGAGCTGCCCTGTGAGATCAGTGTGACTTCTGAGGCTC 1513
Db ATGAATAGGAGGCGTCTGAGAGCTGCCCTGTGAGATCAGTGTGACTTCTGAGGCTC 957
QY CAGTCCAGTCTCAGGCCCATGTGTCAATGGCCAGTGAATAGAGCCCTCAGCTCTGTGTTG 1573
Db CAGTCCAGTCTCAGGCCCATGTGTCAATGGCCAGTGAATAGAGCCCTCAGCTCTGTGTTG 958
QY GTCCTTATTTCT-CCCCATGAGGGCGTGAAGTGTGAGTGTGAGCGGTATATCAATGATGACA 1532
Db GTCCTTATTTCT-CCCCATGAGGGCGTGAAGTGTGAGTGTGAGCGGTATATCAATGATGACA 1017
QY GCTTCTTGACAGGAAGTAGTGTCAAGAAACAGCAGGGGCTTGGCAAGATGATCTAAC 1692
Db GCTTCTTGACAGGAAGTAGTGTCAAGAAACAGCAGGGGCTTGGCAAGATGATCTAAC 1077
QY TGCAAATTCCTACTGGCTCAGCCACAGCAAGTGTCTGTGATCTTGTGAACAAGTTTTTTCAC 1196
Db TGCAAATTCCTACTGGCTCAGCCACAGCAAGTGTCTGTGATCTTGTGAACAAGTTTTTTCAC 1137
QY TTCTCTGAGGCGATCCCTTGGCTACAACAACACAGTGTGTTGTCAGAGATGAATGAG-A 1811
Db TTCTCTGAGGCGATCCCTTGGCTACAACAACACAGTGTGTTGTCAGAGATGAATGAG-A 1197
QY TTCTCTGAGGCGATCCCTTGGCTACAACAACACAGTGTGTTGTCAGAGATGAATGAG-A 1256
Db TTCTCTGAGGCGATCCCTTGGCTACAACAACACAGTGTGTTGTCAGAGATGAATGAG-A 1812
QY ACTCCTTACACCTGTGAATCCAGCACTTTGGAGGCCAAGGCGGTGTGATGGCTTGAAC 1871
Db ACTCCTTACACCTGTGAATCCAGCACTTTGGAGGCCAAGGCGGTGTGATGGCTTGAAC 1257
QY CTGAGAGGTGACAGATGCGGCGAGTCTCACAAGCCCTGTCTGCTGCGGCGCTCTC 1376
Db CTGAGAGGTGACAGATGCGGCGAGTCTCACAAGCCCTGTCTGCTGCGGCGCTCTC 1317
QY TGCCTGAGGCTCCCACTTGGGTGGCAAGGCGAGGCGGCTCCCTCAGCTTGACAGGAGGTGTG 1991
Db TGCCTGAGGCTCCCACTTGGGTGGCAAGGCGAGGCGGCTCCCTCAGCTTGACAGGAGGTGTG 1377
QY GAGCCCTTTCTGTTGGCTGGCCAAAGGCGAGGCGGCTCCCTCAGCTTGACAGGAGGTGTG 2051
Db GAGCCCTTTCTGTTGGCTGGCCAAAGGCGAGGCGGCTCCCTCAGCTTGACAGGAGGTGTG 1991

Db GAGCCCTTTCTGTTGGCTGGCCAAAGGCGAGGCGGCTCCCTCAGCTTGACAGGAGGTGTG 1437
QY GAGGAGAGGCTCAAGCAGGAACCGGGCTGTCCGACGCGCTTCCGGGCGCAGCTGAGATT 2052
Db GAGGAGAGGCTCAAGCAGGAACCGGGCTGTCCGACGCGCTTCCGGGCGCAGCTGAGATT 1497
QY CCGGGTGGGCTGTGGCTGGCGGGCCCGGCACTCGGAGCAGCGGGCCAGCCCTGCGAGGC 2112
Db CCGGGTGGGCTGTGGCTGGCGGGCCCGGCACTCGGAGCAGCGGGCCAGCCCTGCGAGGC 1557
QY CCCGGCAATGAGAGGCTTGAACCCGGGCGAGCGGCTGTGGAGGAGGTGTACTGGGTCCC 2172
Db CCCGGCAATGAGAGGCTTGAACCCGGGCGAGCGGCTGTGGAGGAGGTGTACTGGGTCCC 1617
QY CAGCAGTGCAGCCCGCGGCGCTGTGCTCGATGATTTCTCACTGGGCGCTTGAACACCT 2232
Db CAGCAGTGCAGCCCGCGGCGCTGTGCTCGATGATTTCTCACTGGGCGCTTGAACACCT 1677
QY CAGCAGTGCAGCCCGCGGCGCTGTGCTCGATGATTTCTCACTGGGCGCTTGAACACCT 2292
Db CAGCAGTGCAGCCCGCGGCGCTGTGCTCGATGATTTCTCACTGGGCGCTTGAACACCT 1737
QY TCCGTGTGGCGCGAGGCTCCCGCAGCAGCACCCTGTCACAGGCGCCAGTCCC 2411
Db TCCGTGTGGCGCGAGGCTCCCGCAGCAGCACCCTGTCACAGGCGCCAGTCCC 1797
QY ATGACACGCAAGGCGTGAAGTGTGGGCGCACGCGACCGGAGTGTGAGCAGCAGCTAC 2412
Db ATGACACGCAAGGCGTGAAGTGTGGGCGCACGCGACCGGAGTGTGAGCAGCAGCTAC 1857
QY CCTGTGAGCCCTGTGTGGAATCAGTGGGTGAAGCCAGCTGGGCTCTGAGTGTGGTGG 2472
Db CCTGTGAGCCCTGTGTGGAATCAGTGGGTGAAGCCAGCTGGGCTCTGAGTGTGGTGG 1917
QY AGACTTGGAGACCTTATGCTATAGCTCAGGGAATCGTAATTAACCAATAGAACCTGT 2532
Db AGACTTGGAGACCTTATGCTATAGCTCAGGGAATCGTAATTAACCAATAGAACCTGT 1977
QY GTCTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTGTATCTAGCTACTGTGATGG 2592
Db GTCTAGCTCAGGCTGTGTGAATGACCAATCCACACTGTGTATCTAGCTACTGTGATGG 2037
QY GCTTGTGAGAGCTTATGTCTAGCTCAGGGAATCGTAATTAACCAATAGAACCTGT 2652
Db GCTTGTGAGAGCTTATGTCTAGCTCAGGGAATCGTAATTAACCAATAGAACCTGT 2097
QY TCTAGCTCAAGTTTGTAAACACACCAATAGCAGCTGTGTAGCTCAGGCTATGTGA 2712
Db TCTAGCTCAAGTTTGTAAACACACCAATAGCAGCTGTGTAGCTCAGGCTATGTGA 2157
QY ATGACCAATCGACAGCTGTGTATCTGCTACTTTCATGGGCATCCGTGTGAAGACAC 2772
Db ATGACCAATCGACAGCTGTGTATCTGCTACTTTCATGGGCATCCGTGTGAAGACAC 2217
QY CAACAGGCTTGTGTGAGCAATTAAGCTTCTATCACCTGTGGGTGAGGTGTGAGTGC 2832
Db CAACAGGCTTGTGTGAGCAATTAAGCTTCTATCACCTGTGGGTGAGGTGTGAGTGC 2277
QY GAAAGAGATCAGCAAGGAGATTAAGGTGTGGGCGGCTTTTATAGATTGGGTAGTA 2892
Db GAAAGAGATCAGCAAGGAGATTAAGGTGTGGGCGGCTTTTATAGATTGGGTAGTA 2337
QY GAAAGAGATCAGCAAGGAGATTAAGGTGTGGGCGGCTTTTATAGATTGGGTAGTA 2952
Db GAAAGAGATCAGCAAGGAGATTAAGGTGTGGGCGGCTTTTATAGATTGGGTAGTA 2397
QY GAGGAAATTTACAGTCAAAAGGGGTTTGTCTCTGCGGCGCAGAGTGGGGGTGCGAAG 3012
Db GAGGAAATTTACAGTCAAAAGGGGTTTGTCTCTGCGGCGCAGAGTGGGGGTGCGAAG 3012
QY GTGCTCAGTGGGCTTCTTTTGAAGCAGATGAGCAGGAAAGGACTTTTCAACAAGTA 3071
Db GTGCTCAGTGGGCTTCTTTTGAAGCAGATGAGCAGGAAAGGACTTTTCAACAAGTA 2457
QY GTGCTCAGTGGGCTTCTTTTGAAGCAGATGAGCAGGAAAGGACTTTTCAACAAGTA 3072
Db GTGCTCAGTGGGCTTCTTTTGAAGCAGATGAGCAGGAAAGGACTTTTCAACAAGTA 3131
QY ATGTATCAATTAAGGCAAGACCGGCTTTTACACTCTTTTGTGTGTGTAATGTATCA 2517
Db ATGTATCAATTAAGGCAAGACCGGCTTTTACACTCTTTTGTGTGTGTAATGTATCA 2576

QY	3132	GTAAAGTTGGGGCAGGGCATATTCACCTCTTTTGTGATCTTTCAGTTACTTTCAGGCCATC.	3191
Db	2577	GTTAAGTTGGGGCAGGGCATATTCACCTCTTTTGTGATCTTTCAGTTACTTTCAGGCCATC	2636
QY	3192	TGGGGCTATATGTGCAGATTCACAGGGGATGGGATGGGCTTGGGCTCGGGGCTCAGAGGCTTGA	3251
Db	2637	TGGGGCTATATGTGCAGATTCACAGGGGATGGGATGGGCTTGGGCTCGAGAGGCTTGA	2696
QY	3252	CAGCTACTCTGGTGGGGCTTGGAGATGTTTGTGTGCACACTGTGTATCTAGTTAAATCT	3311
Db	2697	CAGCTACTCTGGTGGGGCTTGGAGATGTTTGTGTGCACACTGTGTATCTAGTTAAATCT	2756
QY	3312	AGTGGGGACGTGGAGAAACCTTTGTGTCTAGCTCAGGAGTTTAAACGACCAATCAGCC	3371
Db	2757	AGTGGGGACGTGGAGAAACCTTTGTGTCTAGCTCAGGAGTTTAAACGACCAATCAGCC	2816
QY	3372	CCTGTCAAAAGACACACTCGGCTCTACCATCAGAGATGTGGTGGGGCCAGATTAAG	3431
Db	2817	CCTGTCAAAAGACACACTCGGCTCTACCATCAGAGATGTGGTGGGGCCAGATTAAG	2876
QY	3432	AGATTAAGACAGGCTGCCGAGCCAGCAGTGGCAACGCGACAGTCCCTATCCACAT	3491
Db	2877	AGATTAAGACAGGCTGCCGAGCCAGCAGTGGCAACGCGACAGTCCCTATCCACAT	2936
QY	3492	ATGGAGAGTTTGTCTTTTGTGTTTGGTGGATTAATCTGTGTACAGTCGTCCGTTTGGGTC	3551
Db	2937	ATGGAGAGTTTGTCTTTTGTGTTTGGTGGATTAATCTGTGTACAGTCGTCCGTTTGGGTC	2996
QY	3552	CACACTGCTTTATGAGCTGTTAACACTCACACGAAAGGTCGACAGTTCACTCCCTGAAGC	3611
Db	2997	CACACTGCTTTATGAGCTGTTAACACTCACACGAAAGGTCGACAGTTCACTCCCTGAAGC	3056
QY	3612	CACTTAAGCACAGAGCCACCGGGAGGAATTAACAATCCGGCGCGCTGCTTAAGAGC	3671
Db	3057	CACTTAAGCACAGAGCCACCGGGAGGAATTAACAATCCGGCGCGCTGCTTAAGAGC	3116
QY	3672	TATAACACTCACCGGAGAGTCTCAGCTTACTCTCAGCAGCAGAGACACAGACCCA	3731
Db	3117	TATAACACTCACCGGAGAGTCTCAGCTTACTCTCAGCAGCAGAGACACAGACCCA	3176
QY	3732	CCAGAGGAGAAACTGCGAACAATCTGAACATCAGAAAGAACAAATCCAGATGACCC	3791
Db	3177	CCAGAGGAGAAAACTGCGAACAATCTGAACATCAGAAAGAACAAATCCAGATGACCC	3236
QY	3792	ACCTTAAGAGGTAAACACTCACGTCGAGGGTCCGGGCTTCCTTGAATGATGAGAG	3851
Db	3237	ACCTTAAGAGGTAAACACTCACGTCGAGGGTCCGGGCTTCCTTGAATGATGAGAG	3296
QY	3852	ACCAAGCACTACAGAGTTTCGAGACAAGCCAGAGATTTAGATTCAGCCGCGCAACAT	3911
Db	3297	ACCAAGCACTACAGAGTTTCGAGACAAGCCAGAGATTTAGATTCAGCCGCGCAACAT	3356
QY	3912	GATGAATATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGGAGCATGTGCT	3971
Db	3357	GATGAATATGCCCTCTCTGCAAAAAAATTTACAAAAATTTGGCGGAGCATGTGCT	3416
QY	3972	CGGTGCTGTGTCCAGCTACGCGGGGAGGTTAAGTGGGAGATTCGCTTAGCCTGGGA	4031
Db	3417	CGGTGCTGTGTCCAGCTACGCGGGGAGGTTAAGTGGGAGATTCGCTTAGCCTGGGA	3476
QY	4032	GGTGAAGACTCAGTAGACTGTGATTTGACACAGCCCTTAGGCTGGGGAGACAGACTGA	4091
Db	3477	GGTGAAGACTCAGTAGACTGTGATTTGACACAGCCCTTAGGCTGGGGAGACAGACTGA	3536
QY	4092	GACCTGTTCCTCCCTCCGCAAAAAAATTTGACAAAAAGTGTAAATAAGGTGCTGATATGG	4151
Db	3537	GACCTGTTCCTCCCTCCGCAAAAAAATTTGACAAAAAGTGTAAATAAGGTGCTGATATGG	3596
QY	4152	CTAGGCGGAGTGGGCTCATGCTGTAAATCCAGCACTTTGGGAACCCAGAGCGGGCGGCTC	4211
Db	3597	CTAGGCGGAGTGGGCTCATGCTGTAAATCCAGCACTTTGGGAACCCAGAGCGGGCGGCTC	3656

QY	4212	ACCTAAGCTCAGGAGTGTGAGACCAAGCTGGCCAACTGAGAAAGCCCATCTTCTTAA	4271
Db	3657	ACCTAAGGCTCAGGAGTGTGAGACCAAGCTGGCCAACTGAGAAAGCCCATCTTCTTAA	3716
QY	4272	AAATCAAAATTAGCGGGCTGTGGGGGAGTGGTGGAGCAATGCCGTAAATCCAGCTACT	4331
Db	3717	AAATCAAAATTAGCGGGCTGTGGGGGAGTGGTGGAGCAATGCCGTAAATCCAGCTACT	3776
QY	4332	CAGAGGCTGAGGCGAGGAGAAATCACTTTAACCACAGAGAGCGCGGTTGCAGTGCAGCGAG	4391
Db	3777	CAGAGGCTGAGGCGAGGAGAAATCACTTTAACCACAGAGAGCGCGGTTGCAGTGCAGCGAG	3836
QY	4392	ATCGTGCCATTGCACTCCACCCTCAGCTCCAGCTGGGGCAACAAGACCACAACTGTCTTAA	4451
Db	3837	ATCGTGCCATTGCACTCCACCCTCAGCTCCAGCTGGGGCAACAAGACCACAACTGTCTTAA	3895
QY	4452	AAAAAAAAAAAAAGTGCCTGACATATAAGAGTGGCATGTGAATAATGTGGCAGGCCAA	4511
Db	3896	AAAAAAAAAAAAAGTGCCTGACATATAAGAGTGGCATGTGAATAATGTGGCAGGCCAA	3955
QY	4512	CATGTTTAAAGATGTGAGACTCTGCTCTCCATGGTCCCTGTTAAAAACCCACCCTCAAG	4571
Db	3956	CATGTTTAAAGATGTGAGACTCTGCTCTCCATGGTCCCTGTTAAAAACCCACCCTCAAG	4015
QY	4572	CCAGGTGCAATGAGCTCATGCTATTAATCCACAGACTTTGGAGGCGAGCGGGTGAATC	4631
Db	4016	CCAGGTGCAATGAGCTCATGCTATTAATCCACAGACTTTGGAGGCGAGCGGGTGAATC	4075
QY	4632	ACCTGAGGTCAGGAGTGTGAGACCAAGCGCTGACCAACAAATGTTGAAATCCACCTGTAC	4691
Db	4076	ACCTGAGGTCAGGAGTGTGAGACCAAGCGCTGACCAACAAATGTTGAAATCCACCTGTAC	4135
QY	4692	TAAATATACAAATATAGATGAGCATAGTGTGTGATGCTCTTAATCCCACTACTTGGAG	4751
Db	4136	TAAATATACAAATATAGATGAGCATAGTGTGTGATGCTCTTAATCCCACTACTTGGAG	4195
QY	4752	GCTGAGGCGAGAAATTCATAGAACACAGGAGCGCGAGGTTGATGAGACCGGAGATCGTG	4811
Db	4196	GCTGAGGCGAGAAATTCATAGAACACAGGAGCGCGAGGTTGATGAGACCGGAGATCGTG	4255
QY	4812	CCATTGCACTCCAGGCTGAGCAATGAGGAAATCCATCTCAAAAAACAACAACAAAA	4871
Db	4256	CCATTGCACTCCAGGCTGAGCAATGAGGAAATCCATCTCAAAAAACAACAACAAAA	4315
QY	4872	CCCACTCTCTACTCCACAGGAGCTGGGTATACAGAGCTGGGGCCACATCAAGTCAAGTCTGT	4931
Db	4316	CCCACTCTCTACTCCACAGGAGCTGGGTATACAGAGCTGGGGCCACATCAAGTCAAGTCTGT	4375
QY	4932	AGCCACAGAGCTAAGGCGGAGCTGACGAGACCGCGGACCAAGATTAACAGTGTGAGATCAG	4991
Db	4376	AGCCACAGAGCTAAGGCGGAGCTGACGAGACCGCGGACCAAGATTAACAGTGTGAGATCAG	4435
QY	4992	TGTGTGAATTCAGAGCTCCCTGCTGCAATTTGGTATACCAACAGAGGGGGCCCCAGACACAGA	5051
Db	4436	TGTGTGAATTCAGAGCTCCCTGCTGCAATTTGGTATACCAACAGAGGGGGCCCCAGACAGA	4495
QY	5052	TGGCCCCATCAGTCAACACATCACTTCTCATCCAGAGATGTCTGTGTTTCTTGACAGCT	5111
Db	4496	TGGCCCCATCAGTCAACACATCACTTCTCATCCAGAGATGTCTGTGTTTCTTGACAGCT	4555
QY	5112	GGGGTAAATTAGGACAGAAGGTGAACAGTCTTGGGTGTGTGATCAGACTGCCCCAGGCA	5171
Db	4556	GGGGTAAATTAGGACAGAAGGTGAACAGTCTTGGGTGTGTGATCAGACTGCCCCAGGCA	4615
QY	5172	GGCCTTGTGGCCTGTAGAAAAGCTTCAAGGCTTAAGCGGGGACAGGCTTACGCTGTGA	5231
Db	4616	GGCCTTGTGGCCTGTAGAAAAGCTTCAAGGCTTAAGCGGGGACAGGCTTACGCTGTGA	4675
QY	5232	ATCCACAGACATTTGGGAGCGCGAGCGGGTGGATCAGAGGCTACAGAGATGTGATCCATC	5291
Db	4676	ATCCACAGACATTTGGGAGCGCGAGCGGGTGGATCAGAGGCTACAGAGATGTGATCCATC	4735
QY	5292	CTGGCTTAACAGCGTGAACCCCGTCTCTACTATAAAATATACAAAAAATTGGCCGCGCATGCT	5351

Db 4736 CTGGCTAACAGGTGAACCCCTCTCTACTATAAAATACAAAATTTGGCGGACATGGT 4795
QY 5352 GGGGGGACACTGATGATTCAGCTACTCGGAGGCTGAGGAGAGAGATGGCTGAACCCG 5411
Db 4796 GGGGGGACACTGATGATTCAGCTACTCGGAGGCTGAGGAGAGATGGCTGAACCCG 4855
QY 5412 AGAGGAGAGATTGGCAGTGAAGCCGAGATGCGCCACTGCACCTCCAGCTTGGGAGACAG 5471
Db 4856 AGAGGAGAGATTGGCAGTGAAGCCGAGATGCGCCACTGCACCTCCAGCTTGGGAGACAG 4915
QY 5472 CAAGACTCATCTGGAAAAAGAAAAAGCTTCAAGTGTGAGCCAGAGAGCCAGGCTG 5531
Db 4916 CAAGACTCATCTGGAAAAAGAAAAAGCTTCAAGTGTGAGCCAGAGAGCCAGGCTG 4975
QY 5532 TAATTTGTCACTTACCATGACCTTGGGAGAGGACACTTCCCTTCCGCGCCAGTTGACG 5591
Db 4976 TAATTTGTCACTTACCATGACCTTGGGAGAGGACACTTCCCTTCCGCGCCAGTTGACG 5035
QY 5592 GGTGGAAATCGACTCCAAAGGTCCCTTCAGCATTAAAGCTGATGCTTCAAGATGAGAA 5095
Db 5036 GGTGGAAATCGACTCCAAAGGTCCCTTCAGCATTAAAGCTGATGCTTCAAGATGAGAA 5095
QY 5652 GATGGGGCAGTTTCCCTCTCTCAGCCAGCCGCTGTCACCTTCAAGGTGAATGACACAG 5711
Db 5096 GATGGGGCAGTTTCCCTCTCTCAGCCAGCCGCTGTCACCTTCAAGGTGAATGACACAG 5155
QY 5712 GAAATGACGTGTCCCAATCCCGAGTTCACAAAGCCCTTGGGAGACCTTACTGTCAAGGTCG 5771
Db 5156 GAAATGACGTGTCCCAATCCCGAGTTCACAAAGCCCTTGGGAGACCTTACTGTCAAGGTCG 5215
QY 5772 TGCACAGAGAGTGAAGGTGACGTGAGCCAAATGCTCCAAAGGCTTCTGCTTATTCGGG 5831
Db 5216 TGCACAGAGAGTGAAGGTGACGTGAGCCAAATGCTCCAAAGGCTTCTGCTTATTCGGG 5275
QY 5832 ACAGACATCGGCTTCTGCTGCTTACCGGAGTTCAGGAGGCTTACAGCGAATGAGTCA 5891
Db 5276 ACAGACATCGGCTTCTGCTGCTTACCGGAGTTCAGGAGGCTTACAGCGAATGAGTCA 5335
QY 5892 TGGGGGGGCGGGGGGTTTCTGGGGGAGTTCACAGCTAATCACTTGGGAGAGACAGCT 5951
Db 5336 TGGGGGGGCGGGGGGTTTCTGGGGGAGTTCACAGCTAATCACTTGGGAGAGACAGCT 5395
QY 5952 GGAACCTTGGAGTGTGCTTATCCAAATGAGGTGAGGAGAGAGAGAGAGAGAGAGAG 6011
Db 5396 GGAACCTTGGAGTGTGCTTATCCAAATGAGGTGAGGAGAGAGAGAGAGAGAGAGAG 5455
QY 6012 CTATATCTCAGATGAGGAGTCAAGAGGTCCTCCAGACAGGAGGCTCCGAGAGATTGGGG 6071
Db 5456 CTATATCTCAGATGAGGAGTCAAGAGGTCCTCCAGACAGGAGGCTCCGAGAGATTGGGG 5515
QY 6072 GTAGGAATGGAGCAACAGGCTTCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6131
Db 5516 GTAGGAATGGAGCAACCA-GCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5574
QY 6132 CAGGCTTGAAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6190
Db 5575 CAGGCTTGAAGATCCCAAG 5634
QY 6191 CGAG 6233
Db 5635 C-AG 5676

RESULT 2
US-60-487-610-19215

; Sequence 19215, Application us/60487610

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: HUANG, Hongjin

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,

; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CU001469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19215
; LENGTH: 190672
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(190672)
; OTHER INFORMATION: n - A,T,C or G, or insertion/deletion polymorphism (see Table
US-60-487-610-19215

Query Match 15.1%; Score 942; DB 7; Length 190672;
Best Local Similarity 73.2%; Pred. No. 9,4e-144;
Matches 1473; Conservative 1; Mismatches 341; Indels 197; Gaps 13;

QY 1873 TGAAGGTGACAGCATGCGGAGTCTCAGACGCTCGTTCCGCTCTGCGCGCTCTCT 1932
Db 178820 TGAAGGTGAGAGCGGTGCTGAGAGGCTCTCAGAGCCCTCGCTCTGCGCGCTCTCT 178879
QY 1933 GCGTGGGCTCCCACTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1992
Db 178880 GCGTGGGCT 178939
QY 1993 AGCCCTTCTTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2052
Db 178940 AGCCCTTCTTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 178999
QY 2053 AGGAGAGCTCAAGCAG 2112
Db 179000 AGGAGAGCTCAAGCAG 179059
QY 2113 CGGGTGGGCTGAGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2171
Db 179060 CAGGTGAGGCTGAGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179118
QY 2172 CCGGAGCAATGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2231
Db 179119 CCGGAGCAATGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179178
QY 2232 CAGCAGTGTGAG 2291
Db 179179 CAGCAGTGTGAG 179234
QY 2292 TCCGAGGAG 2346
Db 179235 TCCGAGGAG 179294
QY 2347 TGGGCTCTGCTGAG 2406
Db 179295 TGGGCTCTGCTGAG 179354
QY 2407 GTCCATTCAGACAGCAG 2466
Db 179355 GTCCATTCAGACAGCAG 179413
QY 2467 GCTACACCTGACAGCAG 2526
Db 179414 GCTACACCTGACAGCAG 179473
QY 2527 GGTGAG 2586
Db 179474 GGTGAG 179533
QY 2587 CCGTGTCTAGCTCAG 2646
Db 179534 CCGTGTCTAGCTCAG 179593
QY 2647 ATGGGGCTTGGAG 2706
Db 179593 ATGGGGCTTGGAG 179650

```
Db 179594 GTGGGCGCTTGGAGACCTTATGCTAGCTCAGGATTTGAACATACAGTACGACT 179653
Qy 2707 CTGATCTAGCTCAAGTTTGTAAACACACACACCTCTGTCTACTACGAGTA 2766
Db 179654 CTGTCTAGCTCAGGGTTTGTAGTGCACCAATGACACCTGTACTACTGTCTGG 179713
Qy 2767 TGTAAATGACACACACGCTGTATCTGTACTACTACTCAATGGGACCTGGTGAAG 2826
Db 179714 TGGGGCCCTTGGAAAACTTTTGTCTAGCTCAGGATTTGAATATACCAATCGGACTG 179773
Qy 2827 ACCACCAAAAGCGCTTGTGTGAGCAATAAAGCTTCAATCCTGGGTGAGGTGGCTG 2886
Db 179774 TGTATCTAGCTCAAGTTTGTAAACAC----- 179800
Qy 2887 AGTCCGAAAAGAGAGTACGCAAGGAGATTAAGGTGGGCGCTTTTATAGATTGGCT 2946
Db 179801 -----ACCAATCAGACCCCTGTGTACTCTCAAGTT 179832
Qy 2947 AGTAAAGAAATTTACAGTCAAAAGGGGTTTGTCTCTGGCGGACAGAGTGGGGGTC 3006
Db 179833 TGTGAGTGCACCAATCCACACTCTGTACTAGCTGCTCTGGTGG----- 179876
Qy 3007 GCAAGGTGCTCAGTGGGGTCTTTTGTAGCCAGGATGAGCCAGAAAAAGACTTTCACA 3066
Db 179877 -----GGCCTTGGAAAACTTTATGTCT 179899
Qy 3067 AGTAAATGCTCATTAATTAAGCAAGACCCGCAATTAACACTCTTTGTGTGTGAATGT 3126
Db 179900 AGCTCAGGATTTGAATATATACCAATCGGCACTGTATCTAGCTCAAGGTTTGAACA 179959
Qy 3127 CATCAGTTAAAGTGGGCGAGGCAATATCACTTCTTGTATCTTCTAGTACTTACG 3186
Db 179960 CACCAATCA-----GCACCTGTGTGTAGTCTCAG 179989
Qy 3187 CCAATCTGGGCGCTATATGTGCAAGTTACAGGGGATGCGATGGCTTGGGCTCAGAG 3246
Db 179990 GTTGTGAGTGCACCAATCCA-----CACTCT 180016
Qy 3247 CTGACAGCTACTCTGTGGGGCGCTTGGAGAAATGTTGTGTGCACTCTGTATCTACTT 3306
Db 180017 GTATCTAGCTCTCTGGGGGCGCTTGGAGAACTTTGTGTGTATCTGTATCT-----T 180072
Qy 3307 AATCTAGTGGGAGCTGAGAACTTGTGTCTAGCTCAGGATTTGAAGCAACCAATC 3366
Db 180073 AATCTAGTGGGAGCTGAGAACTTGTGTCTAGCTCAGGATTTGAAGCAACCAATC 180132
Qy 3367 AGCCGCTGTCAAAACAGACACTGTGGCTTACCAATCAGACAGATGGGTGGGCGCAG 3426
Db 180133 AGCACTGTGTCAAAACAGGCGCACTGTGTACCAATCAGACAGATGGGTGGGCGCAG 180192
Qy 3427 ATAGAGAAATAAAGAGGCGTGGCGGAGCGAGTGGCAAGCGGCACAGTCCCTATCC 3486
Db 180193 ATAGAGAAATAAAGAGGCGTGGCGGAGCGAGTGGCAAGCGGCACAGTCCCTATCC 180252
Qy 3487 ACAATATGGACGCTTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3546
Db 180253 ACAATATGGAGCTTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180309
Qy 3547 GGTGTACACACTGCTTTATGAGCTGTAACTACACACAGAAAGTGTGTGTGTGTGTGT 3606
Db 180310 GGTGTACACACTGCTTTATGAGCTGTAACTACACAGAAAGTGTGTGTGTGTGTGTGT 180369
Qy 3607 GAAGCCACTAGACACAGAGCGCACCGGAGAGTAATGAACACTCCGGCGGCGTGTCTTA 3666
Db 180370 GAAGCCAGTGAACACAGAGCGCACCTGGAGAGAAAGTGTGTGTGTGTGTGTGTGT 180429
Qy 3667 AGAGCTATTAACACTCAGCGCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3726
Db 180430 AGAGCTGTAAACACTCAGCGC--AAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180488
Qy 3727 ACCCAGCAGAAAGAAAGAACTGCAACATCTGAATCAGAAAGAAAGAACTCCAGAT 3786
Db 180489 ACCCAGCAGAAAGAAAGAACTCCAGACATCTGTGAATCAGAAAGGAGACAGCTCCAG 180548
```

```
Qy 3787 GCACCACTTAAGAGCTGTAAACACTGAGAGGTCGGGCGTCTCTTGTGAAGTCA 3846
Db 180549 GCGCCACCTTAAGAGCTGTAAACACTGAGAGGTCGGGCGTCTCTTGTGAAGTCA 180608
Qy 3847 GTGAGACCAACAGCACTCAGCAGGATTTGGAGACA 3878
Db 180609 GTGAGACCAACAGCACTCAGCAGGATTTGGAGACA 180640

RESULT 3
US-09-947-914-48/c
; Sequence 48, Application US/09947914
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON
; FILE REFERENCE: C1001298
; CURRENT APPLICATION NUMBER: US/09/947,914
; NUMBER OF SEQ ID NOS: 75
; SEQ ID NO 48
; LENGTH: 1748349
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1748349)
; OTHER INFORMATION: n = A,T,C or G
US-09-947-914-48

Query Match 13.8%; Score 860; DB 5; Length 1748349;
Best Local Similarity 56.6%; Pred. No. 2,1e-130; Inbels 404; Gaps 16;
Matches 2186; Conservative 0; Mismatches 1270;

Qy 71 AAATATGCGCGGCGATGGTGTGCTCAGCTGTGTAAATCCCTGMACTTTGGGACATCAAGGC 130
Db 1132483 AGAATGAGACTGGAGACAGAGTGGCTCAGCGCTGTATCCAGACACTTTGGAGCGTGAAGGC 1132424
Qy 131 AAGTGCATCTTGAAGTGCAGAGATTTCCAGACTACCTCTGGCCAACTGTGAACCTTAT 190
Db 1132423 GGGCGGTCACCTTGAAGGCGAGAGTTTGAACACCTCTGGCCAACTGTGAACCTTGT 1132364
Qy 191 CTCACATA-----AAAAATACAAAATTTAGCCAGGATGGTGGCAGCACTGTATATCCCG 246
Db 1132363 CTACTACTTATTAATAAATACAGAAATTAAGCCAGGCGTGTGGACGCGCTGTACTCCA 1132304
Qy 247 GCTACTCAGAGAGCTGAGGAGAGCAATCACTTGAATCCAGAGGCGGAGGTTGCAATGA 306
Db 1132303 GCTACTTGGAGAGCTGAGGAGGCGCAAGAACCGCTTGAACCAAGAGGAGGTTGCAATGA 1132244
Qy 307 GCTGATGATCACACACTGCACTGCAAGCTGGGTGACAGAGCAAGACTCTATCTCAAAAA 366
Db 1132243 GCCCAGATCTGTGCCCTCTGCACTGCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 1132184
Qy 367 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 426
Db 1132183 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1132124
Qy 427 AGGAGCTGAGTGGAGAGATCACTTGAACCTGGGAGAGCAAGCTACAGTACAGCAAG 486
Db 1132123 TACTAAAGAAAGATGTAGAGCTGTCTTTTGGGTGTGAAGAAAGAAATCACTTTC 1132064
Qy 487 ATCATGCATACACTCAGCGCTGGGCAACAGAGAGAACCTGTCTCAAAAAATATAT 546
Db 1132063 TCTTTCAACTTTAAGTCAAAACAGATATAAAGAAATACAGCAATATATGTCAAAACCA 1132004
Qy 547 AATATATAAAGAAAAAAGAGCTGTGTATATGTCTCTG---GTCAATCATCTACTATG 603
Db 1132003 GCTGGGTCGGGGAAGCTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1131944
Qy 604 TATATAGTTTCAAACTCAAAAGATCCAGATAGTCAATTTTGTAGCTTGTGGCGCTATG 663
```

Db 1131943 GTCGTAATATCCCATTTAGGCCAATTTAAGCTACCAAGTTTAATGACCGCTGACA 1131884
QY 664 GTCCTGTGCACATCACTGTGCCCTGCTTTCTAGCACAAAAGCAGCTATTAACAATACA 723
Db 1131883 AATTCCTGAAATTTTCATTAATTTGGCTCCCTCGACACACACTGAGAGCTGAGCAAGAG 1131824
QY 724 TACATCAATTTTTTTATAGACATCGAGATTTGAATTTTCATATGATTTTTTACATT--TTAT 780
Db 1131823 GTGAGTGGCTTAGAATTTCTTCATGCCCCGTCTTATTTTGGTTGATTAACCTCAGTTGT 1131764
QY 781 AAAATATCTTTTAA-----AATTTTCCCTAACCACTTTTAAAGTGAAGAGCGCG 834
Db 1131763 TAAATTAACCACTAAAAAGTGAATTTTCACTGATATATCACTTTAACAAGCTTTT 1131704
QY 835 CAGCGCGCATGCTGACAGCCTGTAAATTCAGACCTTTGGAGGCTGAGTGGGACATCA 894
Db 1131703 AATATAGTTTCATCTTCTAGTGTGATTTAAACTCAGGTTTATTAATAATATTCGATT 1131644
QY 895 CTGAGATCAACAGTTTCGAGACCGCTGGCCACATAGCAAAACCCCATTTTCTACTAA 954
Db 1131643 TATAGAAATATTTTTAGAAATATTTTAAGTGAATAAGTGAACAGATGACCCCAA 1131584
QY 955 AATTAATAATATAGCTGGGCAATAGTGTGCACACCTGTGATCCAGCTACTGGAGGCT 1014
Db 1131583 TTTGCAATATRAAAA---TGACTGAAAGGATATACCTCAAAATGTGATATTTA-ATT 1131528
QY 1015 GAGGCGAGAAATCGCTTAACCTGGGAAAGCGAGGTTGCAGTGAAGCAACATCAATGCCA 1074
Db 1131527 GTAAAGATATATGCTTATTTCTTCTTTGCTTTTCTATTTTTCAGATTTTGCGA 1131468
QY 1075 CTGCACTCCAGCCTGGGTGACAGAGTGAAGACTTCGCTCAACGAAAAAAAGTGTAA 1134
Db 1131467 ATGTTTACAACTGGGAAAAAAATTAAGCCAAATTAATATATAGTTTGAAGAAAAATATA 1131408
QY 1135 AGCATATCTTAATTCAGTACATCACTGATACATCACTGAGTGTGCTGCT----- 1185
Db 1131407 TAGCATTTAGATTAATGATATAGCTGCTGACAGAGATGAATCAAGATTGACTATTTTAA 1131348
QY 1186 ----CCTGCTGTGAGGCACTACCTGGAAGTAGAGTGTGCTGTGTCACAGAGCATACAT 1240
Db 1131347 TAAGGTAATATCTTTATCTTTATTTTATGATGTTTACTGCTTCAAGATGTGTCAT 1131288
QY 1241 TTCCACATTTACAGACACTACCAAGTTGCCATCCAGAGGTTTTTTTTTTTCAATCTA 1300
Db 1131287 ATGCATTAATCTAATTTTGAATCACTCTTGACGAGGAGACAGGTTTATATCTTCATTA 1131228
QY 1301 CACTCCCCCAGCAACAATGAGAGTTACTCCAGATCTTTTACAAGATGCTTACGCC 1360
Db 1131227 TACAGATGCAAAAACGATGATTAATGATTTGCCCTGATCAATGCGCTAATTAAGCAG 1131168
QY 1361 AGTACCAATGAAAAAGAGAGTGGAGGGAAGCTGCCAGCCCTTCTTAACCATGAGCA 1420
Db 1131167 TTAAGCTGGGACTTTGACCTGGATTTTCGTTTGTGTAACCCCAAGACCTCTTCAAAA 1131108
QY 1421 AATACCTGTAGAGCTTCTGATGCTGGAAGATCAATACGGGGGCTCTGTGAGCCTG 1480
Db 1131107 CAGCTGGGGGAAGCTTCTCTCATGCTCTCTGTTTAAAGAAAGTCTCACTGAAAACTTA 1131048
QY 1481 CCCCCGTGACAGTGTGACTGTGAGCTTCGAGCCTCCATCCAGTCTCA--GCCCATGTC 1538
Db 1131047 GATCTGATTTCTGCAACGGGTCCCTCTCTGCTGCTTGAAGCAATTTTAATGCCCCATCTTG 1130988
QY 1539 ATGGCAGTGAATATAGCCCTCACTCTCTGTTGCTTTATTTCTCCCATGTTGGGCT 1598
Db 1130987 ACTAGAAATCATCGGGTAGACACATGCATTCAGGCATCGATCGTACACAGATGTGCG 1130928
QY 1599 GAAGTCTGATTTGAGCCGTTATTTCAAGATGACAGCTTTCTTGACAGAAAGTGTGCTA 1658
Db 1130927 CTCACGTGGACTTTAGCTTGAATATGATGAGCTGACGTGAAGAACTTGCCAAAGACA 1130868
QY 1659 -----CAGAAACAGCAGGGGCTTGGCAAGATGATCTAATCGCAAACTCA 1703
Db 1130867 GCACCCACACTCTGGGATGAAGAAACAGGACAAATGTTCTTACAAACATCCAGTTTACAACA 1130808

QY 1704 CCTGGCTCAAGCCACAGCTAGTCTGTGATCTTGAACAAGTTTTTCTACTCTCTGAGGC 1763
Db 1130807 TCTTTAAAGGCCCATACAGAGGTTTGGAAATCATGTTTCCCTATCTCTCCCTTAAGAAAGT 1130748
QY 1764 CATCCCTTGGCTCAACACACACCGATTGTTGTGACAGATGAATGACGAAGTCCCTTACAC 1823
Db 1130747 AGGAGTGGCTTTTATGAGAAATGGCCGAAGGAATTAATCTTGGGGAATCTGTCTAAG 1130688
QY 1824 CTGTAATCCAGACTTTTGGAGGCAAGGCGGGTGGATGAGTGTGAGCCTGAGAGGTGAC 1883
Db 1130687 CTGTGCTGGGTCCCGGTGTGTTGGCTTGGGACATCCGCTCTG---CTGAAGGTGAC 1130631
QY 1884 AGCATCCGGCAGTCTCTACAGCCCTGTTGCTGCTGCGGCTCTCTCTGCTGCTGCTGCT 1943
Db 1130630 AGCCTGCTGACATCCCTCAAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130571
QY 1944 CACTTGGTGGCACTTGAAGAGCCCTTCAAGCCCAACCGCTGCACTGTGGGAGCCCTTTCT 2003
Db 1130570 CACTTGGTGGCACTTGAAGAGCCCTTCAAGCCCAACCGCTGCACTGTGGGAGCCCTTTCT 1130511
QY 2004 GGGCTGGCCAAAGCCACAGAGCCGGCTCCCTCAGCTTCAAGGAGGTGTGAGGAGAGCT 2063
Db 1130510 GGGCTGGCCAAAGGCTGAGAGCCCACTCCGACGCTTGCAGGAGGTGTGAGGAGAGGCA 1130451
QY 2064 CAAGCAGAAACCGGGGCTGCGACAGGCGCTTGGGGCCAGCTGGAATCCGGGTGGGCT 2123
Db 1130450 CGAGCGGAAACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130391
QY 2124 GGGCTGG 2183
Db 1130390 GGGCTTATGTTGG 1130332
QY 2184 GAGCTTTAGACACCCGGGGCCAGCGGCTGGGAGGGGTGACTGGGTGCCCCAGAGTGTGCA 2243
Db 1130331 GGGACTTTAGCACTGG 1130272
QY 2244 CCGCGGG 2303
Db 1130271 CCGCGGG 1130216
QY 2304 GGGCTGGGGAGCTGACAGCCCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2358
Db 1130215 GGGCTGAGGACCTGACAGCCCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1130156
QY 2359 CGGCGGAGGCTGCGGAGGAGCAGCAGCCCGGCTGACAGAGGCGCCAGTCCCATGAGCC 2418
Db 1130155 TGCGCCGAGGCTGCGGAGGAGCAGCAGCCCGGCTGACAGAGGCGCCAGTCCCATGAGCC 1130096
QY 2419 ACAGAGGCTGAGAAAGTGGGGGCGACAGGCGACCGGGAGCTGGCAGGCACTTACCCCTGCA 2478
Db 1130095 ACCCAAGGGCTGAGAAAGTGGAGACAGAGGCA--CAGAGCGGAGCAGCAGCTTACCTGCA 1130037
QY 2479 GCGCTGGTGGGAGATCACTGGGTGAAGCAGCTGGGGCTCTGAGTGTGAGTGTGAGTGTGAG 2538
Db 1130036 GCGCGGCTGGGAGATCACTGGGTGAAGCAGCTGGGGCTCTGAGTGTGAGTGTGAGTGT 1129977
QY 2539 GAGAACCTTATGTTAGGTCAGGAGTCGATGTAATTAACCAATAGCAGCCTGTGTCTAGC 2598
Db 1129976 GAGAGCTTTATATCTAGCTCAAGATTTGTAATTAACCAATAGCAGCCTGTGTCTAGC 1129917
QY 2599 TCAGGCTGTGTAATGACCAATCCACACTCTGTATCTAGTACTGTGATGGGCGCTTGG 2658
Db 1129916 TCAGGCTTTGTAGTGCACCAANTCGGCACTGTATCTAGTACTGTGATGGAGCATGG 1129857
QY 2659 AGAAGCTTTATGTTAGTGTGAGGATTTGTAATTAACCAATAGCAGCCTGTGTATCTAGCT 2718
Db 1129856 AGAGTCTTTATATCTAGCTCAGGAGTGTGTAATTAACCAATAGCAGCCTGTGTATCTAGCT 1129797
QY 2719 CAAGGTTTGAACACCAATCAGCAGCCTGTGTCTAGCTCAGGATGTGTAATGAGCAG 2778
Db 1129796 CAAGGTTTGTAGTGCACCAATCAGCAGCTGTGTATCTAGC----- 1129757

QY 2779 AATGACAGTCTGTATCTGGCTACTTTCATGGGATCCGTGTGAAGACCCACCAACAG 2838
Db 1129756 ----- 1129757
QY 2839 GCTTGTGTGACATAAAGCTTCTATCACCCTGGGTGACAGTGGGCTGATCCGAAAAGA 2898
Db 1129756 ----- 1129757
QY 2899 GAGTCAGGAAAGGATTAAGGTGGGCGCTTTTATAGATTGGTAAAGGAAA 2958
Db 1129756 ----- 1129757
QY 2959 ATTACAGTCAAAAGGGGTTGTCTCTGGCGGACGAGATGGGGGTGCGCAAGTGTCA 3018
Db 1129756 ----- 1129757
QY 3019 GTGGGGGTCTTTTGTGACCCAGATGACCCAGAAAAGACTTTCACAAGTAAATGTCA 3078
Db 1129756 ----- 1129757
QY 3079 CAATTAAAGCAAGGACCCGCAATTACACCTCTTTTGTGGTGAATGTCAATCACTTAAGT 3138
Db 1129710 GTAAATACACCAATCGGACCTGTATCTAGCTCAAGGTTTGTAAATACACCAATCA 1129654
QY 3139 TGGGACAGGCAATATTCATCTTCTTTGTGATTTCTCAGTTACTTCAGGCAATCTGGGGGT 3198
Db 1129653 ----- 1129654
QY 3199 ATATGTGCAAGTTACAGGGGATGCGATGCTTGGCTCAGAGGCTTGACAGCTAC 3258
Db 1129653 --GCACCTGTGTATAGCTCAAGGTTTGTGAATGACCAATCAACACTATACGTAGCTGC 1129596
QY 3259 TCTGTGGGGCTTGGGAAATGTGTGTGTCACACCTCTGTATCTAGTAACTATCTAGTGGG 3318
Db 1129585 TCTGTGGGGCTTGGGAAATGTGTGTGTCACACCTCTGTATCTAGTAACTATCTAGTGGG 1129536
QY 3319 ACGTGAGAACCTTTGTGTCTAGCTCAGGGAATTTGAACGACCAATCAGGCGCTGTCA 3378
Db 1129535 ACGTGGAGAACCTTTGTGTCTAGCTCAGGGAATTTGAACGACCAATCAGGCGCTGTCA 1129476
QY 3379 AAACAGACACTCGGCTTACCAATCAGAGAGATGTGGGTGGGCGCAATTAAGAAATTA 3438
Db 1129475 AAACAGGCACTCGGCTTACCAATCAGAGAGATGTGGGTGGGCGCAATTAAGAAATTA 1129416
QY 3439 AAGGAGCTGCGCGGACGAGAGTGGGAAACGCGGAGTCCCTTCCACATATAGGAG 3498
Db 1129415 AAGGAGCTGCGCGGACGAGAGTGGGAAACGCGGAGTCCCTTCCACATATAGGAG 1129356
QY 3499 CTTTGTCTTTTGTGTGTGGATTAATCTTGTCTACTGCTGCTTTTGGGTCCACACTG 3558
Db 1129355 CTTTGTCTTTTGTGTGTGGATTAATCTTGTCTACTGCTGCTTTTGGGTCCACACTG 1129296
QY 3559 CTTTGTATGAGCTGTAACTACACGAGAGTCTGACGCTTCACTCTGAAGCCACTAG 3618
Db 1129295 CTTTGTATGAGCTGTAACTACACGAGAGTCTGACGCTTCACTCTGAAGCCACTAG 1129236
QY 3619 ACCGAGACCCGCGGAGGAAATGAACAATCCGCGCGCTGCTTAAAGACTTAACA 3678
Db 1129235 ACCGAGACCCGCGGAGGAAATGAACAATCCGCGCGCTGCTTAAAGACTTAACA 1129176
QY 3679 CTCACCGGAGGCTGTGAGCTTCTCCTCAGCAGGAGAGACGACGAGCCGACCTTA 3738
Db 1129175 CTCACCGGAGGCTGTGAGCTTCTCCTCAGCAGGAGAGACGAGCCGACCTTA 1129116
QY 3739 GAAGAACTGCGAACAATCTGAACATGAGAAGAAACTCCAGATGCAACACTTA 3798
Db 1129115 GAAGAACTGCGAACAATCTGAACATGAGAAGAAACTCCAGATGCAACACTTA 1129056
QY 3799 GAGGTAAACACTCACTGCGAGGCTCCGCGCTTCTCTTGAAGTCACTGAGACCAAGC 3858
Db 1129055 GAGGTAAACACTCACTGCGAGGCTCCGCGCTTCTCTTGAAGTCACTGAGACCAAG 1128996
QY 3859 ACTCACCAGTTTCGACACA 3878

Db 1128995 ACCCAACAATTCCGACACA 1128976
RESULT 4
US-60-487-610-19434
; Sequence 19434, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CLO01469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 19434
; LENGTH: 30105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19434
Query Match 12.3%; Score 767; DB 7; Length 30105;
Best Local Similarity 68.8%; Pred. No. 1.5e-115;
Matches 1402; Conservative 0; Mismatches 205; Indels 432; Gaps 8;
QY 1866 TTGACCTGAGAGTGTAGACATGCGGAGTCTCAGAGCCCTGCTCGTGGGCGC 1925
Db 3445 TTGATTAATGACAGGTGACAGCGGTGCTGGCAGTCTCAGAGCCCTGCTTGTGGAC 3504
QY 1926 CTCCTCTGCTTGGCTCCCACTTCTGCTGAGGCACTTGAAGCCCTTACGCGCTGCA 1985
Db 3505 CTCCTCTGCTTGGCTCCCACTTCTGCTGAGGCACTTGAAGCCCTTACGCGCTGCA 3564
QY 1986 CTGTGGAGGCCCTTCTGCTGGGTGGCCAAAGGCGAGAGCGGCTCCCTCAAGTTGAGGA 2045
Db 3565 CTGTGGAGGCCCTTCTGCTGGGTGGCCAAAGGCGAGAGCGGCTCCCTCAAGTTGAGGA 3624
QY 2046 GGTGTGAGAGGAGGCTTCAMAGCAAGAACCGGGGCTGCGACAGGCGCTTGGGCGAGCT 2105
Db 3625 GGTGTGAGAGGAGGAGGCGGACGCGGAAACCGGCGCTGCGTGGGCGCTTGGGCGAGCT 3684
QY 2106 GGAGTTCGGGCTGGGCGTGTG---GCTTGGCGGGGCGCGCACTCGAGACGCGGCGAGC 2162
Db 3685 GGAGTTCGGGCTGGGCGGCGGCGTGTGCGGCGCGAGCACTCGAGACGCGGCGAGC 3744
QY 2163 CTGCGAGGCGCGGCGCAATGAGAGGCTTAAGACCGCGGCGAGCGGCTGCGAGGCTTAC 2222
Db 3745 CTG-CTTGGCGCTTGGCGCAATGAGAGGCTTAAGACCGCGGCGAGGCTGCGAGGCTTAC 3803
QY 2223 TGGGTGCCCCAGACAGTGGCCAGGCGCGCGGCTGTGCTGCTGATTTCTCACTGGGCT 2282
Db 3804 TGGGTGCCCCAGACAGTGGCCAGGCGCGCGGCTGTG---CGCTGATTTCTCACTGAGGCT 3859
QY 2283 TAGCAGCCTTCCCGGCGGAGGCTGCGGACCTGCAAGCCGCAATGCTGAGCTGCC- 2341
Db 3860 TAGCAGCCTTCCCGGAGGAGGCGGCTGCGGACCTGCAAGCCGCAATGCTGAGCTGCCA 3919
QY 2342 ---CTCAGTGGGCTGTGTGGGCGCGAGGCTCCCGAGGAGCAACCCCTGCTCCA 2397
Db 3920 CCCACTCATGGGCTGTGTGGGCGCGAGGCTCCCGAGGAGCAACCCCTGCTCCA 3979
QY 2398 CAGCGCCAGTCCATCAGACAGCAAGGAGGCTGAGAACTGGGCGGCGAGCGGCGGAG 2457
Db 3980 TGGCGCCAGTCCATCAGACAGCAAGGAGGCTGAGAACTGGGCGGAGCGGCGAGC- 4038
QY 2458 TGGCAGGAGCTACCCCTGCAAGCCCTGTGTGGGAAATCACTGTGGTGAAGCCAGTGGGCT 2517
Db 4039 TGGCAGGAGCTACCCCTGCAAGCCCTGTGTGGGAAATCACTGTGGTGAAGCCAGTGGGCT 4098
QY 2518 COTGAGTGTGGTGGAGACTTGGAGAACTTTATGTAGCTCAGGAGATCGTAAATACACC 2577

```
Db 4099 TCTAGCTCTAGTGGGAGCTGAGAGCTTTATGTCTAGCTCAGGGATGTAAATACACC 4158
QY 2578 AATCAGACCCCTGTGTCTAGCTCAGGGTCTGTGAATGCACCAATCCACTCTATCTA 2637
Db 4159 AATCAGACCCCTGTGTCTAGCTCAGGGTCTGTGAATGCACCAATCCACTCTATCTA 4218
QY 2638 GCTACTGTGATGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATGTAAATACACA 2697
Db 4219 GCTGCTGTGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGGATGTAAATACACA 4278
QY 2698 ATCGGCACTGTATCTAGCTCAGGATTTGTAAACACA-CCATCAGACCCCTGTCTA 2756
Db 4279 ATCGGCACTGTATCTAGCTCAGGATTTGTAAATACACCATCCGACCTCTATCTA 4338
QY 2757 GCTCAGGGTATGTGATGACCAATGACAGCTGTATCTGTCTACTCTTCACTGGCAGCC 2816
Db 4339 GCTCAGGGTATGTGATGACCAATGACAGCTGTATCTGTCTACTCTTCACTGGCAGCC 4398
QY 2817 GTGTGAAGAGCCACCAACAGGCTTGTGTGACATTAAGCTTCTATCCTGGGTGC 2876
Db 4399 ACCAATGACACCTGTGTATC----- 4418
QY 2877 AGGTGGCTGAGTCCGAAAAGAGATCAGCGAAGGAGATTAAGGCTGGGCGCTTTATA 2936
Db 4419 ----- 4418
QY 2937 GGATTTGGTAGGTAAGAAATAATACAGTCAAAAGGGGTTTGTCTCTGGCGGAGGA 2996
Db 4419 ----- 4418
QY 2997 GTGGGGGGTGCAGAGTGTCTAGTGGGGGTCTTTTGTAGCCAGATGAGCCAGAAAG 3056
Db 4419 ----- 4418
QY 3057 GACTTTCACAAGTAATGATCATCAATTAAGGCAAGACCCGCCATTTACACCTCTTGT 3116
Db 4419 ----- 4418
QY 3117 GGTGAATGTATCATGTTAAGTTGGGCGAGGGCATATTCATCTTTTGTGATTCCTCAG 3176
Db 4419 ----- 4418
QY 3177 TTACTCAGGCCATCTGGGGGTATATGTGCAAGTTACAGGGATGCATGGCTTGGCTTG 3236
Db 4419 ----- 4418
QY 3237 GGCTCAGAGCCTTGACAGCTACTGTGGGGCTTGGAGAAATGTTGTGCGACTCT 3296
Db 4419 -----TAGCTGCTGTGGGGCTTGGAGAACTGTGTCCAACCTCT 4463
QY 3297 GTATCTAGTTAATCTAGTGGGAGCTGGAGAACCTTTGTGTAGCTCAGGGATTTGTA 3356
Db 4464 GTATCTAATTAATGTGATGGGAGGTGGAGAACCTTTGTATCAAGCTCAGGGATTTGTA 4523
QY 3357 CGACCAATAGGGCCCTGTCAAAAGACAGACACCGGCTTACCAATACACAGATGTGG 3416
Db 4524 CGACCAATAGTGCCTGTGAAACAGCGCACCTGGCTTACCAATACACAGATGTGG 4583
QY 3417 GTGGGGGAGATAGAGAAATAAAGAGCTGCCCGAGCAGAGTGGCAACGGCGCAG 3476
Db 4584 GTGGGGGTAGATTAAGAAATAAAGAGCGGTGCTGGAGCCAGCATTTGGCAACCGCTCG 4663
QY 3477 GTCCCTATCCACAATATGCGAGCTTTGTCTTTGTCTGTTTGGGATTAATCTTGTACTG 3536
Db 4664 GTCCCTTTCCGAGTGTGGGTGCTTTGTCTTTGTGCTTTTGCATTAATCTTGTACTG 4703
QY 3537 CTGGCTTTTGGTCCACACTGCTTTTATAGCTGTAACTCAGCAGCAAGGTCTGAG 3596
Db 4704 CTAGCTC--TGGGTCCACGCTGCTTTTATGAGCTGTAACTCAGCAGCAAGATCTGAG 4761
QY 3597 CTTCACCTCTGAAGCACTTAAGACCAAGCAGCCACGGGAGGATGAACAACTCCGGCCG 3656
|||||
```

```
Db 4762 TTCTACTCTCTGAGCCACGCCAGCAGAGCCCACTGGGAGGAACAACAATCTCAGATG 4821
QY 3657 CGCTGCCCTTAAGACTATAACACTACCGCGCAAGGTGTGACGTTCATCTCTCAGCCAGC 3716
Db 4822 CGCGGCCCTTAAGACTGTAAACACTACCGCGCAAGGTGTGACGTTCATCTCTCAGCCAGC 4881
QY 3717 GAGACCAAGAACCCACCAAGAAAGAACTGCGACACATCTGAACATCAGAAAGACA 3776
Db 4882 GATACTACCAACCCACCAAGAAAGAACTGCGACACATCAGAAATCAGAAAGACA 4941
QY 3777 AACTCAGATGCACACCTTAAGAGCTGTAACTCACTGCGAGGGTCCGGGCTTCTT 3836
Db 4942 GAATTCAGACCGCGCACCTTAAGAGCTGTAACTCACTGCGAGGGTCCGGGCTTCTT 5001
QY 3837 CTGAAGTCACTGAGACCAACCACTACCACTTTCGAGACACCAACCCAGAGATTGAGA 3895
Db 5002 CTGAAGTCACTGAGACCAACCACTACCACTTTCGAGACCAATTAATATGCTTTTGA 5060

RESULT 5
US-60-485-450-11978
; Sequence 11978, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-xung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11978
; LENGTH: 30105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-11978

Query Match 12.3%; Score 767; DB 7; Length 30105;
Best Local Similarity 68.8%; Pred. No. 1,56-115;
Matches 1402; Conservative 0; Mismatches 205; Indels 432; Gaps 8;

QY 1866 TTGAGCTGTAGAGGTGACAGCATGCCGCACTCTCAACAGCCCTCTCGCTCGGCGC 1925
Db 3445 TTGATTAATGACAGGTGACAGCGTCTGCAAGTCTCAGAGCCCTCGCTTGTGGCAC 3504
QY 1926 CTCTCTGCTGGGCTGCCACTTGGGTGCGACTTGAGAGACCTTTCAGCCCGCTGCA 1985
Db 3505 CTCTCTGCTGGGCTGCCACTTGGGTGCGACTTGAGAGACCTTTCAGCCCGCTGCA 3564
QY 1986 CTGTGGAGAGCCCTTCTGTGGGCTGGCCAAAGCCAGACCGGCTCCCTCAGATTGCAAGGA 2045
Db 3565 CTGTGGAGAGCCCTTCTGTAGCTGTGCGCAAGAGCCAGACCGGCTCCCTCAGATTGCAAGGA 3624
QY 2046 GTGTGGAGAGAGAGCTCAAGCAAGAACCGGGGCTGCGCAGCGGCTTCCGGGCGAGCT 2105
Db 3625 GTGTGGAGAGAGAGCGCCAGCGGGAACCGGGCTGCGAGCGGCTTCCGGGCGAGCT 3684
QY 2106 GGAATTCGGGGTGGGCGT---GGCTTGGCGGGGCGCGCACTGTGGAAGACCGGCGAGCC 2162
Db 3685 GGAATTCGGGGTGGGCGGCGGTGGGCTTGGCGGGCGCGCACTGTGGAAGACCGGCGAGCC 3744
QY 2163 CTGCGAAGCCCGCGCAATGAGAGCTTACACCCGGGCGCAAGGGCTGGGAGGTTGATC 2222
Db 3745 CTG-TTGGCCCTGTGGCAATGAGAGCTTACACCCGGGCGCAAGGGCTGGGAGGTTGATC 3803
QY 2223 TGGGTGCGCCAGCAGTGCACAGCCCGCGGCTGTGCTCTGATTTCTCACTGGGCGCT 2282
Db 3804 TGGGTGCGCCAGCAGTGCACAGCCCGCGGCTG---CGCTTATTTCTCACAGAGCT 3859
QY 2283 TAGCAGCCTTCCCGCGGGGCGAGGCTGCGGAGCTGACAGCCCGCATGCTGAGCCTGCC- 2341
|||||
```

```

Db      3860 TAGTGGCCGTTCCCGAGGCGAGGGGCTCGGACCTGAGCCCGGCATGCTGAGCCCTCCCA 3919
QY      2342 ----CTCATGGGCTCTCTGTGGGCGGAGGCTCCCCGACGACACACCCCTGCTCA 2397
Db      3920 CCCACTGATGGGCTCTCTGTGGGCGGAGGCTCCCCGACGACACACCCCTGCTCA 3979
QY      2398 CAGCGGCCGATGCCATCGACGACGAGGCGGTGAGAGAGTGGCGGCGGACGAGCCGGAC 2457
Db      3980 TGGCGGCCGATGCCATCGACGACGAGGCGGTGAGAGAGTGGCGGCGGACGAGCCGGAC 4038
QY      2458 TGGGAGGACGATACCCCTGACGACCCCTGAGGGAATCCAGTGGTGAACGACGCTGGGT 2517
Db      4039 TGGGAGGATGCTCCACCTGACGACCCCTGAGGGAATCCAGTGGTGAACGACGCTGGGT 4098
QY      2518 CCTGAGCTGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTAAATACACC 2577
Db      4099 TCTGAGTCTAGTGGGACCTGGAGAGTCTTTATGTCTAGCTCAGGAGTGTAAATACACC 4158
QY      2578 AATGACACCCCTGTCTAGCTCAGGAGTGTGTGAATGACCAATCCACCTGTATCTA 2637
Db      4159 AATGACACCCCTGTCTAGCTCAGGAGTGTGTGAATGACCAATGACCACTGTATCTA 4218
QY      2638 GCTACTGTGATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTAAATACACA 2697
Db      4219 GCTGCTGTGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTAAATACACA 4278
QY      2698 ATGCGACTCTGTATCTAGCTCAGGAGTGTAAATACACA -CCATACGACACCCCTGTCTA 2756
Db      4279 ATGCGACTCTGTATCTAGCTCAGGAGTGTAAATACACA -CCATACGACACCCCTGTCTA 4338
QY      2757 GCTCAGGATGTGATGACCAATCGACAGTGTGTATGTCTAGCTCAGGAGTGTAAATAC 2816
Db      4339 GCTCAGGATGTGATGACCAATCGACAGTGTGTATGTCTAGCTCAGGAGTGTAAATAC 4398
QY      2817 GTGTGAGAGACCAACCAACAGGCTTGTGTGAGCAATTAAGCTTATCAGCTGGGTGC 2876
Db      4399 ACCAATCGACACTGTGTATC----- 4418
QY      2877 AGGTGGGCTGAGTCCGAAAAAGAGAGTACCGAAGGAGATTAAGGGTGGGCGCTTTTATA 2936
Db      4419 ----- 4418
QY      2937 GATTTGGGTAGTAAAGAAAAATTACAGTCAAGGGGTTTGTCTCTGGCGGAGAGA 2996
Db      4419 ----- 4418
QY      2997 GTGGGGGTGCGAAGGTGCTCAGTGGGGGTGCTTTTGTAGCCAGATGAGCAGAAAG 3056
Db      4419 ----- 4418
QY      3057 GACTTTCACAGGATATGTCATCAATTAAGCAAGAGACCGGCATTTACACTCTTTTGT 3116
Db      4419 ----- 4418
QY      3117 GGTGAGATGTCATCAATTAAGTGGGCGAGGCGATATTCATCTTTTGTATCTTACG 3176
Db      4419 ----- 4418
QY      3177 TTACTTCAGGCGCTGTGGCGTATATGTCAAGTATACAGGGAGTGCATGGCTGGCTTG 3236
Db      4419 ----- 4418
QY      3237 GGCTAGAGAGCTTGACAGCTACTGTGTGGGCGCTTGGAGAAATGTTGTGTGACACTCT 3296
Db      4419 -----TAGCTGTCTGTGGGCGCTTGGAGAACTGTGTGTGCTCAACTCT 4463
QY      3297 GTATCTAGTAACTAGTGGGAGTGGAGAACTTTGTGTAGTCTCAGGAGTGTAA 3356
Db      4466 GTATCTAGTAACTAGTGGGAGTGGAGAACTTTGTGTAGTCTCAGGAGTGTAA 4523
QY      3357 CGCACCAATCAGCGGCTGTCAAAACAGACACTGCGCTTACCAATCGCAGAGATGTGG 3416

```

```

Db      4524 CGCACCAATCAGTGGCCCTGACAAACAGGCGACTGGGCTCTACCAATCAGACAGATGTGG 4583
QY      3417 GTGGGCGCATATAGGAATTAAGACAGGCTGCCGACGACGAGTGGCAAGCGGACAG 3476
Db      4584 GTGGGCGCATATAGGAATTAAGACAGGCTGCCGACGAGTGGCAAGCGGCTGGC 4643
QY      3477 GTCCCATTCACAAATATGACAGTGTGTCTTTGTGTGTGGATTAATCTTGTACAG 3536
Db      4644 GTCCCTTCCGAGTGTGGGTGCTTTGTCTTTGTCTTTGTCAATTAATCTTGTACAG 4703
QY      3537 CTGCGCTTTTGGGTCACACTGTCTTTATGAGCTGTAAACACTCACCAGCAAGTCTGAG 3596
Db      4704 CTGACTC--TGGGTCCACGCTGTTTATGAGCTGTAAACACTCACCAGCAAGTCTGAG 4761
QY      3597 CTTCACCTGTGAAGCAGTAAAGACGACGAGCCACCGGAGAGTAAGCAACTCCGCGC 3656
Db      4762 TTTCACCTGTGAAGCAGTAAAGACGACGAGCCACCGGAGAGTAAGCAACTCCGAG 4821
QY      3657 CGGTGCTTAAAGCTATTAACCTCAGCGGAGAGTGTGAGTGTGACCTTCACTCCAGCAG 3716
Db      4822 CGCGGCTTAAAGCTATTAACCTCAGCGGAGAGTGTGAGTGTGACCTTCACTCCAGCAG 4881
QY      3717 GAGACCAAGACCCACAGAGAGAGAACTGCGAAGTGTGAACATCAGACAGAGAGA 3776
Db      4882 GATACCAAGACCCACAGAGAGAGAACTGCGAAGTGTGAACATCAGACAGAGAGA 4941
QY      3777 AACTCAGATGACACACTTAAAGCTGTAAACACTGACGAGGCTCCGCGCTTCTT 3836
Db      4942 GATACCAAGAGCGGCGACCTTAAAGCTGTAAACACTGACGAGGCTCCGCGCTTCTT 5001
QY      3837 CTGGAAGTGTGAGAGCAAGGACGACGAGTGTGGAGACAGCCAGAGAGTGTAGA 3895
Db      5002 CTGGAAGTGTGAGAGCAAGGACGACGAGTGTGGAGACAGCCAGAGAGTGTAGA 5060

RESULT 6
US-60-487-610-19424/C
; Sequence 19424, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19424
; LENGTH: 108815
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19424

Query Match      12.1%; Score 752; DB 7; Length 108815;
Best Local Similarity 68.8%; Pred. No. 4,6e-113;
Matches 1408; Conservative 1; Mismatches 131; Indels 508; Gaps 7;

```

D 59650 AGGAGAGGTGCGACGCGGAAACGGCCCTGCGTCCGGGCTTGGGGCCAGCTGAGTTC 59591
Q 2113 CGGGTGGGCGTGGGCTTGGGCGGCCCGCACTCGAGACAGCGGCGACCCCTGCGAGCC 2172
D 59590 CGTGTGGGCAATGGGCTTGGTGGCGCAGAGTCCGAGAGACAGGCTGGCCGTG - CTGGCC 59532
Q 2173 CCGGGCAATGAGAGCTTAGACCCGCGGCGCGGCTGCGGAGGCTGACTGGGTGCC 2232
D 59531 CCGGGCAATGAGAGCTTAGACCGGCGCGGCGCGGCTGCGGAGGCTGACTGGGTGCC 59472
Q 2233 AGCAGTCCAGCCCGCGGCGGCTGCTGCTGATTTCTCACTGGGCTTAGGACCTT 2292
D 59471 AGCAGTCCAGCCCGCAGCGAGCTGCTGCTGATTTCTGCGGCGGCTTAGGCTT 59416
Q 2293 CCGCGGGGCGAGGCTGCGGAGCTGCAAGCCCGCATGCTTGAAGCTT 2347
D 59415 CCGCGGGGCGAGGCTGCGGAGCTGCAAGCCCGCATGCTTGAAGCTT 59356
Q 2348 GGGCTCGTGGGCGCGGAGCTGCGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 2407
D 59355 GGGCTCGTGGGCGCGGAGCTGCGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 59296
Q 2408 TCCCATGAGACCGCAGAGGCTGAGAGTGGGCGGCGAGCGAGCGAGCGAGCGAG 2467
D 59295 TCCCATGAGACCGCAGAGGCTGAGAGTGGGCGGCGAGCGAGCGAGCGAGCGAG 59237
Q 2468 CTACCCCTGAGAGGCTGCGGAGCTGCAAGCTGCGGAGCGAGCGAGCGAGCGAG 2527
D 59236 CTACCCCTGAGAGGCTGCGGAGCTGCAAGCTGCGGAGCGAGCGAGCGAGCGAG 59177
Q 2528 GTGAGAGCTTGGAGAGCTTATGCTAGCTCAGGAGTCTGAAATACCAATGAGAGCG 2587
D 59176 GTGAGAGCTTGGAGAGCTTATGCTAGCTCAGGAGTCTGAAATACCAATGAGAGCG 59117
Q 2588 CTGCTGTAGCTCAGGCTGCTGAGTGAATGCAACCAATGCAACTCTG 2631
D 59116 CTGCTGTAGCTCAGGCTGCTGAGTGAATGCAACCAATGCAACTCTG 59057
Q 2632 TATCTAGCTACTGCTGATGGGCGCTTGGAGAGCTTATG 2670
D 59056 TGTGAATGCAACCAATGCAACTGCTGAGTGAATGCAACTGCTGAGAGCG 58997
Q 2671 TCTAGCTCAGGAGTGTAAATACACCAATGCGACTGCTGATGCTAGCTCAGGTTTGA 2730
D 58996 TCTAGCTCAGGAGTGTAAATACACCAATGCGACTGCTGATGCTAGCTCAGGTTTGA 58937
Q 2731 ACACACCAATGAGAGCTGCTGAGTGAATGCAACCAATGCAACTCTG 2790
D 58936 ATACACCAATGAGAGCTGCTGAGTGAATGCAACCAATGCAACTCTG 58877
Q 2791 GTATCTGGCTACTTTCATGGGCAATCCGCTGTAAGAGACCAACCAAGCTTTGTGAG 2850
D 58876 ATCT 58873
Q 2851 CAATAAAGCTTCTATCACTGCTGAGTGAATGCGAGTCCGAAAGAGAGTCAAGAG 2910
D 58872 58873
Q 2911 GGAGATTAAGGTTGGGCGCTTTATAGGATTTGGGTAGGTAAAGAAATTAACAGTCAA 2970
D 58872 58873
Q 2971 GGGGGTTTGTCTCTGCGCGGCGAGAGTGGGGGCTCCAGAGTCTCAGTGGGGGTGCTT 3030
D 58872 58873
Q 3031 TTTGAGGAGATGAGGAGGAGAAAGACTTTTCAAGAGTATGTCATCAATTAAGCAA 3090
D 58872 58873
Q 3091 GGAGCCGCAATTTACACCTCTTTTGTGATGATGTCATGATTAAGTTGGGCGAGGCA 3150
D 58872 58873

Q 3151 TATTCACCTTCTTTGTGATCTTCACTTACCTTACAGGCCATCTGGCGTATATGTCAAGT 3210
D 58872 58873
Q 3211 TACAGGAGATCGATGGCTTGGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 3270
D 58872 58873
Q 3271 TTTGAGATTTTGTGTCAGACTGCTGATCTATGATTAATGAGGAGAGTGAAGAGC 3330
D 58854 TTTGAGATTTTGTGTCAGACTGCTGATCTATGATTAATGAGGAGAGTGAAGAGC 58797
Q 3331 TTTGTCCTAGCTCAGGAGTGTAAAGCAGCAATCAAGCGCTTGTCAAAACAGACACT 3390
D 58796 TTTGTCCTAGCTCAGGAGTGTAAAGCAGCAATCAAGCGCTTGTCAAAACAGACACT 58737
Q 3391 CCGGCTTACCAATGAGAGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3450
D 58736 CAGCTTACCAATGAGAGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 58677
Q 3451 CGAGCAGCAGTGGCAACGCGCAGAGGCTCCTATTCACAAATATGAGAGCTTGTCTTTT 3510
D 58676 GGAGCAGCAGTGGCAACGCGCAGAGGCTCCTATTCACAAATATGAGAGCTTGTCTTTT 58617
Q 3511 GCTGTTTGCATTAATCTTGTCTACTGCTGCTTTTGGGTCACACTGCTTTATGAGCT 3570
D 58616 GCTGTTTGCATTAATCTTGTCTACTGCTGCTTTTGGGTCACACTGCTTTATGAGCT 58557
Q 3571 GTACACCTCAGCAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3630
D 58556 GTACACCTCAGCAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 58497
Q 3631 CCGGAGGAGTGAATCAACTCCGCGCGGCTGCTTGAAGCTATTAACACTGACCGGAG 3690
D 58496 CCGGAGGAGTGAATCAACTCCGCGCGGCTGCTTGAAGCTATTAACACTGACCGGAG 58437
Q 3691 GTCTGAGCTTCACTCTGACAGCAGGAGCAGCAACCAACCAACCAACCAACCAACCA 3750
D 58436 GTCTGAGCTTCACTCTGACAGCAGGAGCAGCAACCAACCAACCAACCAACCAACCA 58377
Q 3751 AACACATCTGAATCAAGTGAAGAGCAACCACTCAGATGAGCAGCAGCTTGAAGCTTGA 3810
D 58376 AACACATCTGAATCAAGTGAAGAGCAACCACTCAGATGAGCAGCAGCTTGAAGCTTGA 58317
Q 3811 TCACCTGAGGAGTCCGCGGCTTCTTGAAGTCAAGTGAAGAGCAGCAGCAGCAGCTT 3870
D 58316 TCACCTGAGGAGTCCGCGGCTTCTTGAAGTCAAGTGAAGAGCAGCAGCAGCAGCTT 58257
Q 3871 CCGACACA 3878
D 58256 CCGACACA 58249

RESULT 7
US-60-485-450-11975/c
; Sequence 11975, Application us/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: C1001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11975
; LENGTH: 108815
; TYPE: DNA
; ORGANISM: Homo sapiens

US-60-485-450-11975

Query Match 12.1%; Score 752; DB 7; Length 108815;
Best Local Similarity 68.8%; Pred. No. 4,6e-113;
Matches 1408; Conservative 1; Mismatches 131; Indels 508; Gaps 7;

QY 1873 TGAGAGTGAACAGATCCCGGACATCTCCACAGCCCTGCTGCTGCGGCGCTCT 1932
DB 59830 TGAAGGTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59771
QY 1933 GCGTGGGCTCCCACTTGGTGGGACCTTGAAGAGCCCTTCAAGCCCAACCGCTCAGTGGG 1992
DB 59770 GCGTGGGCTCCCACTTGGTGGGACCTTGAAGAGCCCTTCAAGCCCAACCGCTCAGTGGG 59711
QY 1993 AGCCCCCTTGGGCTGCGCCAGAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2052
DB 59710 AGCCCCCTTGGGCTGCGCCAGAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 59651
QY 2053 AGGAGAGGCTCAAGAGAGAGCCGCGGCTGCGGAGCCGCTGCGGAGCCAGCTGGAGTTC 2112
DB 59650 AGGAGAGGCTCAAGAGAGAGCCGCGGCTGCGGAGCCGCTGCGGAGCCAGCTGGAGTTC 59591
QY 2113 CGGGTGGGCGTGGGCTTGGCGGGGCCCGCACTCGGAGCGGCGGAGCCCTGCGGAGGCC 2172
DB 59590 CGGTGGGCGATGGGCTTGGTGGGCGGAGCACTCGGAGCGAGAGCGGCTGGCGCTG-CTGGCC 59532
QY 2173 CCGGGCAATGAGAGGCTTACAGCCCGGCGGAGCGGCTGCGGAGGCTTACTGGGCTCCC 2232
DB 59531 CCGGGCAATGAGAGGCTTACAGCCCGGCGGAGCGGCTGCGGAGGCTTACTGGGCTCCC 59472
QY 2233 AGCAGTGCACGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
DB 59471 AGCAGTGCACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59416
QY 2293 CCGCGGGGCGAGGCGCTGCGGAGCTGCAAGCCCGCATGCTGAGCT-----CCCTCCAT 2347
DB 59415 CCGCGGGGCGAGGCGCTGCGGAGCTGCAAGCCCGCATGCTGAGCT-----CCCTCCAT 59336
QY 2348 GGGCTCTGCTGGGCGCGGAGCTTCCCGAGAGCAACACCCCTGCTCCACAGGCGCCAG 2407
DB 59336 GGGCTCTGCTGGGCGCGGAGCTTCCCGAGAGCAACACCCCTGCTCCACAGGCGCCAG 59296
QY 2408 TCCCATGAGACCAAGAGGCTGAGAAAGTGGGCGGAGCGGAGCGGAGCGGAGCGAG 2467
DB 59296 TCCCATGAGACCAAGAGGCTGAGAAAGTGGGCGGAGCGGAGCGGAGCGGAGCGAG 59237
QY 2468 CTACCCCTGCAAGCCCTGCTGCGGAACTCCAGTGGTGAAGCCAGCTGGGCTCTGAGTCTG 2527
DB 59236 CTACCCCTGCAAGCCCTGCTGCGGAACTCCAGTGGTGAAGCCAGCTGGGCTCTGAGTCTG 59177
QY 2528 GTGAGACTGGAGAACCTTATGCTAGCTCAGGAGATCGTAATATACCAATACAGCAC 2587
DB 59176 GTGAGACTGGAGAACCTTATGCTAGCTCAGGAGATCGTAATATACCAATACAGCAC 59117
QY 2588 CTGCTGTAGCTCAGGCTGCTGATGAGCAACCAATCCACACTCTG----- 2631
DB 59116 CTGCTGTAGCTCAGGCTGCTGATGAGCAACCAATCCACACTCTG----- 59057
QY 2632 -----TATCTAGCTACTGTGATGGGCTTGGAGAACCTTTATG 2670
DB 59056 TGTGATATGACCAATCGACATATCTAGCTACTGTGATGGGCTTGGAGAACCTTTATG 58997
QY 2671 TCTAGCTCAGGAGATTTGTAATATACCAATCGGCACTGCTGCTATGCTCAAGCTTTGTA 2730
DB 58996 TCTAGCTCAGGAGATTTGTAATATACCAATCGGCACTGCTGCTATGCTCAAGCTTTGTA 58937
QY 2731 ACACCAATATGAGACCTGCTGATGAGTATGATGATGATGATGATGATGATGATGATG 2790
DB 58936 ATACCAATATGAGACCTGCTGATGAGTATGATGATGATGATGATGATGATGATGATG 58877
QY 2791 GTATCTGGCTACTTTCATGGGCAATCCGTGTAAGAGACCAACACAGGCTTTGTGTAG 2850
DB 58876 ATCT----- 58873

QY 2851 CAATAAGCTTCTATCAGCTGGGCTGAGGTGGGCTGAGTCCGAAAAAGAGTCAAGCAAG 2910
DB 58872 ----- 58873
QY 2911 GGAGATTAAGGCTGGGCGCTTTTATAGATTTGGTAGTAAGGAAATATACAGTCAA 2970
DB 58872 ----- 58873
QY 2971 GGGGCTTTGCTCTGGCGGCGAGAGTGGGGGGTCCGAAAGTCTCAGTGGGGTCTT 3030
DB 58872 ----- 58873
QY 3031 TTTGAGCCAGATGAGCAAGGAAAGACATTTCAACAAAGTATGTCATCAATTAAGCAA 3090
DB 58872 ----- 58873
QY 3091 GGACCGGCAATTAACCTTTTGTGGTGAATGTCAATGATTAAGTTGGGCGAGGCA 3150
DB 58872 ----- 58873
QY 3151 TATTCATCTTTTGTGATCTTCACTTCAAGGCAATCTGGGCGTATATGTCAAAT 3210
DB 58872 ----- 58873
QY 3211 TACAGGAGATGAGTGGCTTGGCTTGGCTGAGAGGCTTGAACAGTACTGCTGGTGGCC 3270
DB 58872 ----- 58873
QY 3271 TTTGAGAAATTTTGTGAGACATCTGATATGATTAATCTAGTGGGAGCGGAGAAC 3330
DB 58872 ----- 58873
QY 3331 TTTGAGAAATTTTGTGAGACATCTGATATGATTAATCTAGTGGGAGCGGAGAAC 3390
DB 58872 ----- 58873
QY 3391 TTTGAGAAATTTTGTGAGACATCTGATATGATTAATCTAGTGGGAGCGGAGAAC 3450
DB 58872 ----- 58873
QY 3451 TTTGAGAAATTTTGTGAGACATCTGATATGATTAATCTAGTGGGAGCGGAGAAC 3510
DB 58872 ----- 58873
QY 3511 GCTGTTGCGATTAATCTTGTCTAGTCTGCTGCTTGTGGTCCACACTGCTTTATGACT 3570
DB 58872 ----- 58873
QY 3571 GTCACACTCAGCAAGAGTCTGCAAGCTTCACTCTCTGTAAGCCAGTGAAGCCAGCAAG 3630
DB 58872 ----- 58873
QY 3631 CCGGAGAGATTAACCAATCCGCGCGGCTGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 3690
DB 58872 ----- 58873
QY 3691 CCGGAGAGATTAACCAATCCGCGCGGCTGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 3750
DB 58872 ----- 58873
QY 3751 AACACATGTAACATCAACAAAGAAACCAATCTCAATGACCACTTAAGCTGTAACAC 3810
DB 58872 ----- 58873
QY 3811 TCACTGAGAGGCTCGCGGCTTCTTCTTGAAGTCAAGTCAAGCAAGCACTCAAGCTT 3870
DB 58872 ----- 58873
QY 3871 CGGACACA 3878
DB 58872 ----- 58873

OY 3661 GCCTTAAGAGCTATTAACCTCAGCCGGAAGGCTGACCTTCACCTCCAGCCAGCGAGA 3720
 Db 7883293 ACCTTAAGAGCCGTACACTCAGCCGGAAGGCTGACCTTCACCTCCAGCCAGCGAGA 7883352
 OY 3721 CCAGAACCCACACGAGAAGAAACTGCGAACAACATGTGAACATCAGAGAAGAACT 3780
 Db 7883353 CCAGAACCCACACGAGAAGAAATTCGAACAACATCCGAACATCAGAGAAGAACT 7883412
 OY 3781 CCAGATGACACACCTTAAGAGCTGTAACTGCGAGGGTCCGCGCTTCCTTTG 3840
 Db 7883413 CCAGACGCGCCACCTTAAGAGCTGTAACTGCGAGGGTCCGCGCTTCCTTTG 7883472
 OY 3841 AAGTCAGTGAACCCAGACGACTCACCAGTTGGGACACA 3878
 Db 7883473 AAGTCAGTGAACCCAGACGACTCACCAGTTGGGACACA 7883510

 RESULT 9
 ; Sequence 50, Application US/10450826
 ; GENERAL INFORMATION:
 ; APPLICANT: JI, Darren
 ; APPLICANT: Axelrod, Douglas W.
 ; APPLICANT: Cook, Jonathan S.
 ; APPLICANT: Jaiswal, Neelam
 ; APPLICANT: Eistein, Richard
 ; APPLICANT: Houghton, Adam
 ; APPLICANT: Mertz, Lawrence
 ; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
 ; FILE REFERENCE: 044921-5039-WO
 ; CURRENT FILING DATE: 2003-06-18
 ; PRIOR FILING DATE: 2003-06-18
 ; PRIOR APPLICATION NUMBER: US 60/255,882
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: US 60/285,691
 ; NUMBER OF SEQ ID NOS: 149
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 121724
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. AL136139
 US-10-450-826-50

 Query Match 11.8%; Score 733.2; DB 6; Length 121724;
 Best Local Similarity 90.1%; Pred. No. 5.1e-110;
 Matches 856; Conservative 0; Mismatches 78; Indels 16; Gaps 6;

 OY 1867 TGAGCCCTGAGAGGTGACAGCAATGCGCGCAGCTCAGACGCCCTGTTGCTGCGGCC 1926
 Db 4692 TGCAATGAGAGGTGACAGCAATGCGCGCAGCTCAGACGCCCTGTTGCTGCGGCC 4751
 OY 1927 TCCTCTGCTGGGCTCCCACTTGGGTGGACACTTGAAGAGCCCTTCAAGCCACCGCTGAC 1986
 Db 4752 TCCTCTGCTGGGCTCCCACTTGGGTGGACACTTGAAGAGCCCTTCAAGCCACCGCTGAC 4811
 OY 1987 TGTGGGAGCCCTTTCGCGGCTGGGCAAGGCAAGCCGCGTCCCTCAGCTTGAAGGAG 2046
 Db 4812 TGTGGGAGCCCTTTCGCGGCTGGGCAAGGCAAGCCGCGTCCCTCAGCTTGAAGGAG 4871
 OY 2047 GTGTGAGGAGAGGCTCAACACAGAAACCGGGCTGCGACAGCGGCTTGCAGGACAGTG 2106
 Db 4872 GTGTGAGGAGAGGCTCAACACAGAAACCGGGCTGCGACAGCGGCTTGCAGGACAGTG 4931
 OY 2107 GAGTTCGCGGCTGGGCTGGGCTGGGCGGCGCCGCTCAGGAGAGCGGCGCAAGCCCTGC 2166
 Db 4932 GAGTTCGCGGCTGGGCTGGGCTGGGCGGCGCCGCTCAGGAGAGCGGCGCAAGCCCTGC 4991
 OY 2167 CAGGCCCCGGGCAATGAGAGGCTTGAACACCGGGCGCAGCGGCTCGGAGAGGTGTACTGG 2226
 Db 4992 C-GGCCCCGGGCAATGAGAGGCTTGAACACCGGGCGCAGCGGCTCGGAGAGGTGTACTGG 5050

QY 2227 TGCCACAGAGTGCAGAGCCCGCGGCGGTGTGTCTGCTTCACATTTTCTACACBGGCCTTAGC 2286

Db 5051 TCCTTCACAGATGCGCAGCCACACAGCCGTG---CGCTCGATATCTGGGCGGCGCTTAGC 5106

QY 2287 AGCCTTCCCGGGGGCAGGGCTCGGGAACCTGCAGCCCGCCATTCGCTGAGCTT-----CCC 2341

Db 5107 TGCCCTTCCCGGGGAGCGGGGGCTCGGAAACCCGACGCCGCAATGCTTGAAGCTTCCACCCC 5166

QY 2342 CTCCATGGGCTCTGTGTGGGCGCCGAGCCTCCCGACAGCACACACCCCTGTCTCAGAGC 2401

Db 5167 CTGCGTGGGCTCTGTGTGGGCGCCAAAGCCTCCCGAGTGAAGCGCCACACCACTGCTCCAGAGC 5226

QY 2402 GCCCAGTCCCATCCAGCCACCGAAGGGCTGAGAAAGTGGGGCGCACGGCACCAGGACTGGC 2461

Db 5227 GCCCAGTCCCATCCAGCCACCGAAGGGCTGAGAAAGTGGGGCGCAAG---CGGACCTAGC 5283

QY 2462 AGGCAGCTACCCCTGCAGCCCTGGGTGGGGAATCCAGTGGGTGAAGCCAGCTGGGCTCTG 2521

Db 5284 AGGCAGCTCCACCTGCAGCCCTGTGCGGGATCCACTAGTGAAGCCAGCTGGGCTCTG 5342

QY 2522 AGTCTGTGTGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCGTAAATACACCAATC 2581

Db 5343 AGTGGGTGGGAGCTGGAGAACATTTATGTCTAGCTCAGGGATCGTAAATACACCAATC 5402

QY 2582 AGCACCTGTGTCTAGCTCAGGGGTCTGGAATGCACCAATCCACACTCTGATCTAGCTA 2641

Db 5403 AGCACCTGTGTCTAGCTCAGGGGTCTGGAATGCACCAATCCACACTCTGATCTAGCTA 5460

QY 2642 CTCTGATGGGCGCTGGAGAACCTTTATGTCTAGCTCAGGGATCGTAAATACACCAATC 2701

Db 5461 CTCTGATGGGCGCTGGAGAACCTTTATGTCTAGCTCAGGGATCGTAAATACACCAATC 5520

QY 2702 GCACCTGTGTCTAGCTCAGGGATTTGTAACACACCAATCCAGCCCTGTGTCTAGCTCA 2761

Db 5521 GCACCTGTGTCTAGCTCAGGGATTTGTAACACACCAATCCAGCCCTGTGTCTAGCTCA 5580

QY 2762 GGGTATGTGAATGCACCAATGCAGACGTGTGATCTGGCTACTTTTACAGGG 2811

Db 5581 GGGTATGTGAATGCACCAATGCAGACGTGTGATCTGGCTACTTTTACAGGG 5630

RESULT 10

US-60-487-610-19466/C

Sequence 19466, Application US/60487610

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: HUANG, Hongjin

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,

TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001469

CURRENT APPLICATION NUMBER: US/60/487,610

CURRENT FILING DATE: 2003-07-17

NUMBER OF SEQ ID NOS: 97101

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19466

LENGTH: 29405

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc. feature

LOCATION: (1)---(29405)

OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables

US-60-487-610-19466

Query Match 11.6%; Score 720.6; DB 7; Length 29405;

Best Local Similarity 68.0%; Pred. No. 4,8e-108;

Matches 1395; Conservative 0; Mismatches 129; Indels 527; Gaps 7;

QY 1873 TGAAGAGTGACAGATCCGCGAGTCTCTACAGCCCTGCTGCTTCGGCGCTCTCT 1932

Db 26948 TGAAGAGTGACAGTGTGTGGCACTCTCTACAGCCCTGCTGCTTCGGGGAATCTCT 26889

QY 1933 GCGTGGGCTCCACTTCGGTGCGACTGAGAGACCCCTTACAGCCACCGCTGCACTGTGG 1992
 DB 26888 GCGTGGGCTCCACTTCGGTGCGACTGAGAGACCCCTTACAGCCACCGCTGCACTGTGG 26829
 QY 1993 AGCCCTTTCTGGGCTGGCCCAAGGCCAGAGCCGGCTCCCTGACCTTGGAGAGGTGG 2052
 DB 26828 AGCCCTTTCTGGGCTGGCCCAAGGCCAGAGCCGGCTCCCTGACCTTGGAGAGGTGG 26769
 QY 2053 AGGAGAGGCTCAGCAGAGAACCGGGCTGCGCAGCGCGCTTGGGGCCAGCTGGAGTTG 2112
 DB 26768 AGGAGAGGCTCAGCAGAGAACCGGGCTGCGCAGCGCGCTTGGGGCCAGCTGGAGTTG 26709
 QY 2113 CGGGTGGGCTGGGCTTGGCGGGCCCGCAGCTCGAGAGCAGGGCCAGCCCTCCAGGCC 2172
 DB 26708 CGGGTGGGCTGGGCTTGGCGGGCCCGCAGCTCGAGAGCAGGGCCAGCCCTCCAGGCC 26650
 QY 2173 CCGGGCAATGAGAGGCTTAGCACCAGGGCCAGGGCTGGGGAGGTGATAGGTGCTCC 2232
 DB 26649 CCGGGCAATGAGAGGCTTAGCACCAGGGCCAGGGCTGGGGAGGTGATAGGTGCTCC 26590
 QY 2233 AGCAGTGCAGAGCCCGCGGCTGTGCTGCTGATTTCTACCTGGGCTTACAGACCTT 2292
 DB 26589 AGCAGTGCAGAGCCCGCGGCTGTGCTGCTGATTTCTACCTGGGCTTACAGACCTT 26554
 QY 2293 CCGCGGGGCGAGGGCTCGGAGCTGCAAGCCCGCATGCTGAGCTCC-----CTCCAT 2347
 DB 26553 CCGCGGGGCGAGGGCTCGGAGCTGCAAGCCCGCATGCTGAGCTCC-----CTCCAT 26494
 QY 2348 GGGCTCTGTGGCGGGCCGAGCCCTCCCGAGCAGACACACCCCTGCTCCACAGCGCC 2407
 DB 26493 GGGCTCTGTGGCGGGCCGAGCCCTCCCGAGCAGACACACCCCTGCTCCACAGCGCC 26434
 QY 2408 TCCCATGCAACGACGAGAGGCTGAGAGTGGCGGCGCAGCGCAGCGGAGCTGCGAGCAG 2467
 DB 26433 TCCCATGCAACGACGAGAGGCTGAGAGTGGCGGCGCAGCG--TGGGAGACTGGCAGGTAG 26375
 QY 2468 CTACCCCTGCGAGCCCTGTGGCGGAATCCACTGGGTGAGAGCAGCTGGGCTCTGAGTCTG 2527
 DB 26374 CTACCCCTGCGAGCCCTGTGGCGGAATCCACTGGGTGAGAGCAGCTGGGCTCTGAGTCTG 26315
 QY 2528 GTGGAGACTTGGAGAACCTTATGTAGCTCAGGAGTCTGTAAT----- 2572
 DB 26314 GTGGAGACTTGGAGAACCTTATGTAGCTCAGGAGTCTGTAAT----- 26255
 QY 2573 -----ACACCAATCAGACCCCTGTGTAGCTCAGGCTG 2608
 DB 26254 CTATCTAGCTCGAGGTTGTAAACACACCAATCAGACCCCTGTGTAGCTCAGGCTTGG 26195
 QY 2609 TGAATGCACCAATCCACACTCTGTATCTAGCTACTGATGGGGCTTGGAGAACCTTTA 2668
 DB 26194 TGAATGCACCAATCCACACTCTGTATCTAGCTACTGATGGGGCTTGGAGAACCTTTCA 26135
 QY 2669 TGTCTAGCTCAGGAGTTGTAATACCAATCGGCACTGTATCTAGCTCAAGTTTGT 2728
 DB 26134 TGTCTAGCTCAGGAGTTGTAATACCAATCGGCACTGTATCTAGCTCAAGTTTGT 26075
 QY 2729 AAACACACCAATGAGACCTGTGTAGCTCAGGAGTATGTGAATCAGCAATCAGACGT 2788
 DB 26074 AAACACACCAATGAGACCTGTGTAGCTCAGGAGTATGTGTAGTACCAACCGACACT 26015
 QY 2789 CTGTATCTGGCTACTTTCATGGGCAATCCGTGTGAAGAGACACCAACAGGCTTTGTG 2848
 DB 26014 CTGTATCT----- 26007
 QY 2849 AGCAATTAAGCTTCTATCACTGGGTGAGGTGGCTGAGTCCGAAAGAGAGTACGCA 2908
 DB 26006 ----- 26007
 QY 2909 AGGAGATTAAGGTGGGCGCTTTTATAGGATTTGGGTAGTAAAGAAAATTACAGTCA 2968
 DB 26006 ----- 26007

QY 2969 AAGGGGTTTGTCTCTGGCGGCGAGAGTGGGGGTCCGAAAGTCTCAGTGGGGTGC 3028
 DB 26006 ----- 26007
 QY 3029 TTTTGGAGCAGATGAGCCAGGAAAAAGACTTTCACAAAGTAAATGATCAATTAAGGC 3088
 DB 26006 ----- 26007
 QY 3089 AAGACCCCGCATTTACACCTTGTGTGGGAATGTATCAGTTAAGTTGGGCAAGG 3148
 DB 26006 ----- 26007
 QY 3149 CATATTCATCTTTTGTGATTTCTGATTTACTTACAGCCATCTGGCGGTATATGTCAA 3208
 DB 26006 ----- 26007
 QY 3209 GTTACAGGGGATGCGATGGCTTGGGCTCAGAGGCTTGAACAGTACTCTGGTGGG 3268
 DB 26006 ----- 26007
 QY 3269 CTTTGGAGATGTGTGTGTCAGACTCTGTATCTAGTTAATCTAGTGGGAGCTGGAGAA 3328
 DB 25990 CTTTGGAGATGTGTGTGTCAGACTCTGTATCTAGTTAATCTAGTGGGAGCTGGAGAA 25931
 QY 3329 CTTTGTGTCTAGCTCAGGAGTTGTAAAGCAGCAATCAGCCCTGTCAAAACAGACA 3388
 DB 25930 CTTTGTGTCTAGCTCAGGAGTTGTAAAGCAGCAATCAGCCCTGTCAAAACAGACA 25871
 QY 3389 CTGCGGCTCTACCAATCAGAGATGTGGGGGGGCGCAGATTAAGAAATTAACAGAGCTG 3448
 DB 25870 CTGCGGCTCTACCAATCAGAGATGTGGGGGGGCGCAGATTAAGAAATTAACAGAGCTG 25811
 QY 3449 CCGGAGCAGACAGTGCACAGCGCAGACAGTCCCTATCTACCAATATGAGAGCTTGTCTT 3508
 DB 25810 CCGGAGCAGACAGTGCACAGCGCAGACAGTCCCTATCTACCAATATGAGAGCTTGTCTT 25751
 QY 3509 TTGCTGTGGCGTAATCTTCTGCTACTGCTGCTTTTGGGTCCACACTGCTTTATAG 3568
 DB 25750 TTGCTGTGGCGTAATCTTCTGCTACTGCTGCTTTTGGGTCCACACTGCTTTATAG 25691
 QY 3569 CTGTAACTACACCAAGAGTGTGAGCTTCACTCTCTGTAAGGCACTTAAGACAGAGCC 3628
 DB 25690 CTGTAACTACACCAAGAGTGTGAGCTTCACTCTCTGTAAGGCACTTAAGACAGAGCC 25631
 QY 3629 CACCGGAGAGATGAACAACTCCGGCGGCTGCTTAAAGAGTATTAACACTCAGCGCA 3688
 DB 25630 CACCGGAGAGATGAACAACTCCGGCGGCTGCTTAAAGAGTATTAACACTCAGCGCA 25571
 QY 3689 AGGTCTGACGCTTCACTCTCAGCAGCGAGCAGCAAGAACCCAGAAAGGAAGAACTG 3748
 DB 25570 AGGTCTGACGCTTCACTCTCAGCAGCGAGCAGCAAGAACCCAGAAAGGAAGAACTG 25511
 QY 3749 CGAACAATCTGAATCAGAAAGAGAACAACTCCAGATGACCC--ACCTTAAGAGCTGTAA 3807
 DB 25510 CGAACAATCTGAATCAGAAAGAGAACAACTCCAGATGACCC--ACCTTAAGAGCTGTAA 25451
 QY 3808 CACTCACTGCGAGAGGTCCGGGCTTCTCTTGAAGTCAAGTGAACCAAGCACTCAGCAG 3867
 DB 25450 CACTCACTGCGAGAGGTCCGGGCTTCTCTTGAAGTCAAGTGAACCAAGCACTCAGCAG 25391
 QY 3868 TTTCGGACACA 3878
 DB 25390 TTTCGGACACA 25380

RESULT 11
 US-60-485-450-11933/C
 ; Sequence 11933: Application us/60485450
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: CHANG, Sheng-Yung
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C

;; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
;; FILE OF INVENTION: THEREOF
;; FILE REFERENCE: CL001470
;; CURRENT APPLICATION NUMBER: US/60/485,450
;; CURRENT FILING DATE: 2003-07-09
;; NUMBER OF SEQ ID NOS: 47859
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 11933
;; LENGTH: 13040
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(13040)
;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-485-450-11933

Query Match 11.4%; Score 710.4; DB 7; Length 13040;
Best Local Similarity 63.3%; Pred. No. 2e-106;
Matches 1425; Conservative 0; Mismatches 606; Indels 221; Gaps 13;

QY 3253 AGCTACTGTGGGGCCTTGGAGAAATGTTGTGCGACACTCTGTATCTAATCTA 3312
DB 10290 AGCTACTGTGGGGCCTTGGAGAAATGTTGTGCGACACTCTGTATCTAATCTA 10231
QY 3313 GTGGGAGCTGGAGAACTTGTGTCTGCTCAGGATTTGTAAGCGCAATCAGCGCC 3372
DB 10230 GTGGGAGCTGGAGAACTTGTGTCTGCTCAGGATTTGTAAGCGCAATCAGCGCC 10171
QY 3373 CTGTCAAAACAGACACTCGGCTTACCAATCAGCAGATGTGGGTGGGCCAGATTA 3432
DB 10170 CTGTCAAAACAGACACTCGGCTTACCAATCAGCAGATGTGGGTGGGCCAGATTA 10111
QY 3433 GAATTAAGCAGAGCTGCCGAGCCAGCAGAGTGGCAAAGCGCAGAGTCCCTATCA 3492
DB 10110 GAATTAAGCAGAGCTGCCGAGCCAGCAGAGTGGCAAAGCGCAGAGTCCCTATCA 10051
QY 3493 TGGCAGCTTGTCTTGTCTGCTTGTGCTGATTAATCTGCTACTGCTTGTGGTTC 3552
DB 10050 TGGCAGCTTGTCTTGTCTGCTTGTGCTGATTAATCTGCTACTGCTTGTGGTTC 9991
QY 3553 ACACGCTTTTATGAGCTGTAACTCACTCAAGAGTCTGCAAGTTCACCTCTGAAGC 3612
DB 9990 ACACGCTTTTATGAGCTGTAACTCACTCAAGAGTCTGCAAGTTCACCTCTGAAGC 9931
QY 3613 ACTAAGACGAGCCGACCGGAGGAAATGAACAACCTCGCGCGCTGCTTAAGAGCT 3672
DB 9930 AGCGAGACGAGCCGACCGGAGGAAATGAACAACCTCGCGCGCTGCTTAAGAGCT 9872
QY 3673 ATAACTACACCGCAAGTGTGCAAGTTCCTCTCAACCGCAGCAGACAGCAAGCCAC 3732
DB 9871 ATAACTACACCGCAAGTGTGCAAGTTCCTCTCAACCGCAGCAGACAGCAAGCCAC 9812
QY 3733 CAGAAGGAAGAACTGCAACATCTGAAATGAAAGAAACAAATCCAGATGACCA 3792
DB 9811 CAGAAGGAAGAACTGCAACATCTGAAATGAAAGAAACAAATCCAGATGACCA 9752
QY 3793 CTTAAAGCTGTAACTCACTGCGAGGTCGCGGCTTCTTTGAAGTCAATGAGA 3852
DB 9751 ACTTAAAGCTGTAACTCACTGCGAGGTCGCGGCTTCTTTGAAGTCAATGAGA 9692
QY 3853 CCAAGCACTCAACAGATTTCGACACAGCCAGAGATTGAGATCAGCTGGGCAATG 3912
DB 9691 CCAAGCACTCAACAGATTTCGACACAGCCAGAGATTGAGATCAGCTGGGCAATG 9636
QY 3913 ATGAATGCTCCTCTGTGCAAAATTTGCGGAGGATGCGTGTGTC 3972
DB 9635 GCGAGA-----CTGTTTCAAAATTTAGTCGACGGCCGCGCTGTGCT 9584
QY 3973 CGTGTGTGTGTCAGCTACGCGGAGGCTTAAGTGGAGAGATGCTTGAAGCTGGAG 4032
DB 9583 CAGGCTGTAAATCCAGACTTTGGAGGCGGAGGTGGCGGATCATTGAGAGTGAAG 9524

QY 4033 GTGAAGACTGAGTAGCTGTGATTTTACACAGCCCTAGAGCTGGGAGACAGACTAG 4092
DB 9523 TTGCGAGA-----CCACCTGGCCACATGCTGA 9495
QY 4093 ACCCTGTTCCTCCCGCAAAATTTGACAAAGTGAATGAAGAGTCCGTGAATGCG 4152
DB 9494 ACTGTCTCTACTATTAATTAACA-----GAC 9469
QY 4153 TAGCGGAGTGGCTATGCTGTAAATCCAGACTTTGGAGCCGAGGCGGGGTCA 4212
DB 9468 AGGCAAGGTAACTACAGCTTGTGATCCAGACTTTGGAGCCGAGGCGGGGTGAT 9409
QY 4213 CCTAAGTTCAGAGTGTGAGACAGCAGCTGGCCAAATGAGAAAGCCATCTCTTAA 4272
DB 9408 CACAAGTTCAGAGTGTGAGACAGCAGCTGGCCAAATGAGAAAGCCATCTCTTAA 9357
QY 4273 AATACAAATTAAGCGGCTGTGGGGAGTGTGAGCATGCTGTAAATCCAGCTACTC 4332
DB 9356 ACTAAATTAAGCG-----GGTACACAGCAGCTGTAAATCCAGCTACTC 9311
QY 4333 AGGAGCTGAGGAGAGAGATATCACTTGAATCCAGAGAGGCGGCTTGCAGTGAAG 4392
DB 9310 AGGAGCTGAGGAGAGAGATATCACTTGAATCCAGAGAGGCGGCTTGCAGTGAAG 9251
QY 4393 TCGTGCATTTGACTTCACCCAGCTCAGCTGGCAACAAAGAGCCAACTGTCTTAA 4452
DB 9250 TCACACACTGCACTTCAGCC-----TGGCGACAGAGCCAACTGTCTTAA 9201
QY 4453 AAAAAATTAAGTGCCTGACATTAAGAGGTGTGCAATGATTTGCCAGGCAAC 4512
DB 9200 GAAACCAAAACA----- 9188
QY 4513 ATGTTAAGATGTGAGCTCCTGCTTCATGCTCTGTAAATCCAGCTTGAAGC 4572
DB 9187 -----AAACAAATTAAGTTC 9169
QY 4573 CAGGTGAGTGGCTATGCTAATCCAGACTTTGGAGGCGGAGGCTGATGA 4632
DB 9168 CAGGTGAGTGGCTATGCTAATCCAGACTTTGGAGGCGGAGGCTGATGA 9109
QY 4633 CCTGAGTTCAGAGTTCGAGACAGCTGACCAACATGCTGAATCCACTTACT 4692
DB 9108 CTTGAGTGTGAGAGTGTGAGGCTGCGGTGAGC-----GATGATGTGCTGACCTGAC 9056
QY 4693 AAAAAATTAAGTGAAGTGTGATGCTGTAAATCCAGCTTGAAGC 4752
DB 9055 CAGCTGAGTGAAGTGAAGTGTGATGCTGTAAATCCAGCTTGAAGC 8996
QY 4753 CTGAGGAGGAAATCAGTGAAGAGGAGGAGGAGTGTGAGGAGGAGTGTG 4812
DB 8995 GTCACTGAGGAGGAGGAGTGTGAGGAGGAGGAGTGTGAGGAGGAGTGTG 8936
QY 4813 CATTCACCTCCAGCTGAGCAATGAGCAATCCATCTCAAAAAACAAACAA 4872
DB 8935 GCACCACTGAGGAGGAGGAGTGTGAGGAGGAGGAGTGTGAGGAGGAGTGTG 8876
QY 4873 CCACCTCTACTCCAGGAGGAGTGTGAGGAGGAGGAGTGTGAGGAGGAGTGTG 4932
DB 8875 ACTAAATTAAGTGAAGTGTGAGGAGGAGGAGTGTGAGGAGGAGTGTG 8816
QY 4933 GCCACAGAGCTAAGGCGGAGGAGTGTGAGGAGGAGGAGTGTGAGGAGTGTG 4992
DB 8815 GAAAGCTGAAGGCGGAGGAGTGTGAGGAGGAGGAGTGTGAGGAGGAGTGTG 8756
QY 4993 GTGTGAGATCAGAGCTCCCTGCTTGTGAGCAGCAGGAGGAGGAGGAGGAGT 5052
DB 8755 -----GAAAGCTGTCTACTTAATAAACAACAAAAAATGAGTGTGAGT 8707
QY 5053 GGCCTCCTCAGTCAACATCACTTCACTCAGAGATGTCTTCTTGGCAGCTG 5112
DB 8706 GCGGCGACCAATTAATCCAACTACTTGGAGAGCTGAG--GCAGAGAAATTCCTGAAC 8649
QY 5113 GGGTAAATTAAGACAGAAAGTCACTTGGGTGTGCTGAGTCAAGACTGCCCGCAG 5172

Db 8648 CGGAGGTGGAGGTGGAGTGAACCGAGATGGCCGCACTGCATCCACCTGACCGTACAG 8589
QY 5173 GCCTTGTGGCTGTAGAAAACGTTCAAGCCTAGCCGGGACGGTGGCTCAAGCCTGTA 5232
Db 8588 AGTGAAGCTCTATCTCAAAAAAAGAGAGAGCGGTGGCTGCTCAAGCCTGTA 8529
QY 5233 TCCAGAGCTTTGGAGGCGGAGGCGGTGGATCAGAGCTCAGAGATGCTGACATCC 5292
Db 8528 TCCAGAGCTTTGGAGGCGGAGATGAGCGGATCAGAGCTCAGAGATGCTGACATCC 8469
QY 5293 TGGCTAACAGCGTGAACCCCGCTCTACTATAAAATAC---AAAAATTTGGCGGGCATG 5349
Db 8468 TGGCTAACAGCGTGAACCCCGCTCTACTATAAAATACAAAAAATTTAGCCGGCATG 8409
QY 5350 GTGGCGGACCTGTAGTTCCAGTACTCGGAGGCTGAGGACAGAGATGGCGTGAAC 5409
Db 8408 GTGGCGGACCTGTAGTTCCAGTACTCGGAGGCTGAGGACAGAGATGGCGTGAAC 8349
QY 5410 CGAGAGGACAGTTTGGAGTGAAGCCGAGATCCGCCACTGACCTCCAGCTGGCGACAG 5469
Db 8348 CGGAGGTGGAGCTTGAAGTGAAGAGATCGCCACTGCACTCCAGCTGGCGACAA 8289
QY 5470 AGCAAGACTCCATCTGGAAAAAGAAAAAGAAAA 5501
Db 8288 AGCGAGACTCGGTCTCAAAAAAATAATAATA 8257

RESULT 12
US-60-485-450-11934/c
: Sequence 11934, Application US/60485450
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: CHANG, Sheng-yung
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
: TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
: FILE REFERENCE: CL001470
: CURRENT APPLICATION NUMBER: US/60/485,450
: CURRENT FILING DATE: 2003-07-09
: NUMBER OF SEQ ID NOS: 47859
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11934
: LENGTH: 48853
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(48853)
: OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-485-450-11934

Query Match 11.4%; Score 710.4; DB 7; Length 48853;
Best Local Similarity 63.3%; Pred. No. 2.2e-106;
Matches 1425; Conservative 0; Mismatches 606; Indels 221; Gaps 13;

QY 3253 AGCTACTGTGTGGGCGCTTGGAGATGTTGTGTGACACTGTGTATCTAGTTAATCTA 3312
Db 3260 AGCTACTGTGTGGGCGCTTGGAGATGTTGTGTGACACTGTGTATCTAGTTAATCTG 3201
QY 3313 GTGGGAGCGTGAAGACCTTGTGTGTAGCTCAGGGATTGTAAAGCAACCAACAGCCG 3372
Db 3200 GTGGGAGCGTGAAGACCTTGTGTGTAGCTCAGGGATTGTAAAGCAACCAACAGCCG 3141
QY 3373 CTGTCAAAACAGACACTGCGCTCTACATCAGCAGATGTGGTGGGCCAGATAGA 3432
Db 3140 CTGTCAAAACAGACACTGCGCTCTACATCAGCAGATGTGGTGGGCCAGATAGA 3081
QY 3433 GATATAAAGCAGGCTGCCAGACAGAGTGGCAAGCGCACAGGTCCCTATCCACAAATA 3492
Db 3080 GATATAAAGCAGGCTGCCAGACAGAGTGGCAAGCGCACAGGTCCCTATCCACACAG 3021

QY 3493 TGGCAGCTTTGTCTTTTGTCTTGTGGCATTAATCTTGTCTACTGCTGTGGCTTGTGGCTC 3552
Db 3020 TGGAGGCTTTGTCTTTTGTCTTGTGGCATTAATCTTGTCTACTGCTGTGGCTTGTGGCTC 2961
QY 3553 ACAGCTTTTATGAGCTGTAACTCACTGACCAAGAGGCTGACCTTCACTCCGAAAGCC 3612
Db 2960 ACAGCTTTTATGAGCTGTAACTCACTGACCAAGAGGCTGACCTTCACTCCGAAAGCC 2901
QY 3613 ACTAAGACACAGAGCCGAGGAGATGAACTCCGCGCGCTGCTTGAAGAGCT 3672
Db 2900 AGCGAGACACAGAGCCGAGGAGATGAACTCCGCGCGCTGCTTGAAGAGCT 2842
QY 3673 ATTAACACTCACCGGAGAGTGTGACGCTTCACTCTCAACGAGGACACAGAACCCAC 3732
Db 2841 GTAACTACTCGCGGCAAGGTTGAGGCTTCACTCTCAACGAGGACACAGAACCCAC 2782
QY 3733 CAGAAGGAAGAACTGGGAACACATCTGAAATATGAAAGGAACAACTCCAGATGACCA 3792
Db 2781 CAGAAGGAAGAACTGGGAACATATCGAAACATGAAAGGAACAACTCCAGATGACCA 2722
QY 3793 CCTTAAGAGCTGTAACTCACTGAGAGGCTCGCGGCTTCTTGAAGTCAGTGAGA 3852
Db 2721 ACTTAAGAGCTGTAACTCACTGAGAGGCTCGCGGCTTCTTGAAGTCAGTGAGA 2662
QY 3853 CCAAGCACTACACGCTTTCGAGACAAAGCCAGAGATTTGAGATCAGCTGGGCAACATG 3912
Db 2661 CCAAGCACTACCACTTCTGACACA---GTGCACTGCACTCAGCCTGGGTGACAGA 2606
QY 3913 ATGAATGCGCCTCTCTCAAAAAAATAATAATAATAATTTGGCGGACGATGCTGCTC 3972
Db 2605 GCGAGA-----CTTTTCANAAAAAATAATAATACTGACGCGCGGCGGTGCTGCT 2554
QY 3973 CGTGTGTGTGTCACGCTACGCGGAGGCTTAAAGTGGAGATGCTGAGCCTGGGAG 4032
Db 2553 CAGGCTGTATCCACACTTTGGAGGCGGAGGTGGGAGATCACTTGAAGTCACTAG 2494
QY 4033 GTGAAGCTGACAGTGTGATGATTTACCAAGCCCTCTAGGCTGGGGGACAGATGAG 4092
Db 2493 TTGAGA-----CCAGCCTGGCCCAACATGCTGTA 2465
QY 4093 ACCCTGTTTCCCTCCGCAAAAAAATGACAAAAGTGTAAATGAGGTGCTGATATAGGC 4152
Db 2464 ACTGTGTCTACTATAATAATCA-----GAC 2439
QY 4153 TAGCGGAGTGTGCTCATGCTGTAAATCCAGCACTTTGGAGCCGAGCGGCGGTCA 4212
Db 2438 AGGACAGGTAACTCACTGAGGCTTGTGATCCAGCACTTTGGGACAGCAAGGGGAGTGTAT 2379
QY 4213 CTTAAGTCAAGAGTGTGAACCAAGCTGGCCAACTGAGAAAGCCCATCTCTTAA 4272
Db 2378 CACAAAGTCAAGAGTGTGGGACAGCTGACCAACATGTTGAAACTGCTCT----- 2327
QY 4273 AATACAAATTTAGCGGCGCTGTGGGAGAGTGTGAGCATGCGGTAAATCCAGCTATC 4332
Db 2326 ACTTAAATAATTAAGCG-----GTTGACACAGCACTGTATCCAGCCAGTCT 2281
QY 4333 AGGAGCTGAAGAGAGATCACTTGAACCCAGAGGCGGCGGCTGTGACATGAGCGAGA 4392
Db 2280 AGGAGCTGAAGAGAGATTCGCTTGAATCAAGAGGTGTGACTGTGACATGAGCGAGA 2221
QY 4393 TCGTGCATTTGCACTCAACCACTCAAGCTGGGCAACAAGAGCCAACTCTCTTAA 4452
Db 2220 TCACACCACACTCCAGCC-----TGGGGAGAGAGCCGAAATCCATCTCAAA 2171
QY 4453 AAAAAAAGTGTGACATATTAAGAGTGTGCAATGCAATAGTTGCCAGGCAAC 4512
Db 2170 GAAACACAAACA----- 2158
QY 4513 ATGTTTAAGATGTGAGCTCTGCTTCCATGCTCTGTTAAAAACCAACCTCAAGGC 4572
Db 2157 -----AAACAATTAATAAATTAAGC 2139
QY 4573 CAGGTGACGTGCTCATGCTTAATATCCAGCACTTTGGAGGCGGAGGCGGTGATCA 4632

Db 2138 CAGGTGTGTACACGACCTGTATCCACCTCTCGTGGAGGCGGAGGATCA 2079
Qy 4633 CCTAGGTGAGGATTTGAGACACGCTGACCCACACATGCTGAATCCACCTTACT 4692
Db 2078 CTTAGGTGTGAGGATGAGGCTCGGTGAGC-----GGTGAATTTGGCCACTGCACTC 2026
Qy 4693 AAAAAATCAAAATTAAGATGAGCATGTGTGTCATGCTGTAAATCCACCTTGGGAGG 4752
Db 2025 CAGCTAGCTAGATGATGACCTTGAGAAAAGTAATATTTTAAATAAAGACTGG 1966
Qy 4753 CTGAGGAGGAAAATTCATAGAACAGGAGGCGGAGGTTAGTAGACCGAGATCTGTC 4812
Db 1965 GTCAAGTACAGGCCCAAGTGTTCACACCTGTATCCAGATTTTGGAGCCCGAGGACAGC 1906
Qy 4813 CATGTCACTCCAGCTGAGCATAGAGCAAACTCCATCTCAAAAAAACACACAAAAAC 4872
Db 1905 GGACCACTTGAGGAGGAGGATTCAGAACACACCTGGCCACATGTATGAACCCCATCTCT 1846
Qy 4873 CCACTCTTACTCCAGGAGGAGTGTACAGAGTGGGCGACATCAGTCAAGGTGTGA 4932
Db 1845 ACTAAAAATACAAAAATTAAGCCCGGCTGCGGGGCTGACACTTGTAACTTACCTTGG 1786
Qy 4933 GCCACAGAGCTAAGCGGAGCTGACAGACCGGACAGATTAAGTGTGTGAGATCACT 4992
Db 1785 GAAGGTGTGAAGCGGCGGATCATGAGTCAGGAATTCAGACTAGCTTGGCCAAATACAGT 1726
Qy 4993 GTGTGAGATCAGAGTCCCTCCCATGTGTGACACAGAGGAGGCGCCCAAGACAGAT 5052
Db 1725 -----GAAACGCTGTCTCTACTAAAAACAAAAAAAGTTGTGTGGGTGAGGT 1677
Qy 5053 GGCCCATCCAGTACACCATCCACTTCTCATCCAGAGATGTCTGTTCCTGGACGCTG 5112
Db 1676 GCGCGGACCACTTATATCCCACTTCTGGAGGCTGAG--GCAGAGATTTGCTTGAAC 1619
Qy 5113 GGGTAATTTAGACAGAGATGTGACAGTCTTGGGTGTGTCAGTCAAGTGTGCCAGGCA 5172
Db 1618 CGGAGAGTGTGAGTGTGACAGTGAACCGAGATGCGGCACCTGACCTGAGAGGACAG 1559
Qy 5173 GCGCTTGTGGCTGTAGAAAACGTTCAAGCCCGGCGGACGCTGTGCTCAACGCTGTAA 5232
Db 1558 AGTAGACTTATCTCAAAAAAATAAAGCAAGGAGTGGGTGTCTACAGCTGTCTTAA 1499
Qy 5233 TCCAGCACTTTGGAGGCGGAGGCGGTGATCAGAGTGTGAGAGATGTGACCATTC 5292
Db 1498 TCCAGCACTTTGGAGGCGGAGGCGGTGATCAGAGTGTGAGAGATGTGACCATTC 1439
Qy 5293 TGGCTAACAGCGTGAACCCCGTCTCTACTAAAAATAC--AAAAATTTGGCGGCGATG 5349
Db 1438 TGGCTAACAGCGTGAACCCCGTCTCTACTAAAAATACAAAAAATTTACCGGCGATG 1379
Qy 5350 GTGGCGGCGACCTTACTGCTAGTCTCTGAGGAGGCTGAGGAGAGATGCGGTGAAC 5409
Db 1378 GTGGCGGCGACCTTACTGCTAGTCTCTGAGGAGGCTGAGGAGATGCGGTGAAC 1319
Qy 5410 CGAGAGGAGAGTTTGGAGTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5469
Db 1318 CGGAGGAGTGTGAGTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1259
Qy 5470 AGCAAGACTCCATCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5501
Db 1258 AGCAGAGACTCGGTCTCAAAAAAATAAATA 1227

RESULT 13
US-60-487-610-20017
; Sequence 20017, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARILLI, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20017
; LENGTH: 40207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-20017

Query Match 11.1%; Score 693; DB 7; Length 40207;
Best Local Similarity 65.3%; Pred. No. 1.4e-103;
Matches 128; Conservative 0; Mismatches 580; Indels 127; Gaps 16;

Qy 848 TCAGCCGTGTAATTCACGACCTTTGGAGGCTGAGGCTGGGAGATCACTTGAATCAACA 907
Db 2125 TCACGCGTGTATCCACACTTTGGAGGCGCAAGTGGCGGATCACA--AGTCAAGA 2182
Qy 908 GTTGAGACAGCCCTGGGCAACATAGCAAAACCCCATTTCTACTAAAAATAAATAAATA 967
Db 2183 GTTCAAGACACGCTTGACCAACATAGTGAACCCCGTCTTACGAAAGTACAAAAATTA 2242
Qy 968 GCTGGCAATGATGTGACACACCTGTGATCCAGTACTTGGAGGCTGAGGAGAGAT 1027
Db 2243 GCCAGGCAATGATGTGTGACACCCGTAATCCAGTACTGAGAGGCTGAGGAGAT 2302
Qy 1028 CGCTTGAACCTGGGAGGAGGAGTGTGACGTGACACATCATGCTGCTGCTGCTGCTG 1083
Db 2303 CGCTTGAACCTGGGAGGAGGAGTGTGACCGGCGGAGGCTGCTGCTGCTGCTGCTG 2362
Qy 1084 AGCCGGGTGACAGAGTGAAGTGTGCTGCAACGAAAAAAGTGTAAAGGCAATTC 1143
Db 2363 AGCCGGGTGACAGGCGGAGGAGTGTGCTGCAACGAAAAAAGTGTAAAGGCAATTC 2419
Qy 1144 TAATCACTGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
Db 2420 TCCCTTTTGTGTTAGAGAGGCAAAACATTAATGATGATGATGATGATGATGATGATGAT 2479
Qy 1198 CATACCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
Db 2480 TAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2539
Qy 1258 ACTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
Db 2540 ATGCTGTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2599
Qy 1318 AATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
Db 2600 CAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659
Qy 1378 GG-----AAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1416
Db 2660 AGTCTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2719
Qy 1417 AAGAAATTCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1473
Db 2720 ATGATTAATGCAAAAGAGTGTGCACTAATTAATTAATTAATTAATTAATTAATTAAT 2779
Qy 1474 GAGGCTG-----CCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1509
Db 2780 GAGCATGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2839
Qy 1510 CCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1564
Db 2840 ATACATTTGGGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2899
Qy 1565 CTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1624
Db 2900 CTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2959
Qy 1625 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1684

Db 2960 CTAGACAGCAAAATACATATTTAATGAGGCAAAATGCTCCGAACCCCTGAGAGAC 3019
Qy 1685 GATCTAATCTGCAAAATCTCTACTGCTGCTCAGCCACAGACTTCTGTGATCTTGACAGT 1744
Db 3020 TTGGGAGGGCTTTTGAAGTTTAAAGAAAGATAGTAAGAGAGATTTCATTGAAAAAG 3079
Qy 1745 TTTTTCATCTTCT--CTGAGGCCATCCCTTGGCTACACACACCAGTGGTTGACAGATG 1802
Db 3080 AATAGAAATGTAGTTGAAGCACTTCTCAGAGTGTAGTCCAAAGCTTTCCTGTAGAGAG 3139
Qy 1803 AATAGCAAGTCCCTTACACCTGTAAATCCAGACACTTGGGAGGCCAGGGGGGAGAT 1862
Db 3140 GTGAGGCACAGAGACAGACATATGACCTACTCTGTATAGTTGTGTGACTGTGTA- 3198
Qy 1863 GGCCTGAGCCTGAGAGGTGACAGCATCCGCGAGTCCCTCACAAGCCCTTCTGCTCTCG 1922
Db 3199 ---GTCAATGTGAGAGATGACAGCGTGTGAGAGTCTCACAAGCCCTGCTGCTCTCG 3255
Qy 1923 GCGCTCTCTGCTGCTGCTGCTCCACTTGGTGGCACTTGAAGAGCCCTTACAGCCACGCT 1982
Db 3256 GGCATCTCTGCTGCTGCTGCTCCACTTGGCGCACTTGAAGAGCCCTTACAGCCACGCT 3315
Qy 1983 GCACTGTGGAGCCCTTCTGCGGCTGGCGAAGCGCGGCTCCCTCAGCTTGGCAG 2042
Db 3316 GCACTGTGGAGCCCTTCTGCGGCTGGCGAAGCGCGGCTCCCTCAGCTTGGCAG 3375
Qy 2043 GAGGTGTGAGAGGAGAGGCTCAAGAGAGACCGGGGCTGCGCAGCGGCTTGGCGGCA 2102
Db 3376 GGAAGGGGTGAGAGAGGCGGAGCGGAGACCGGGGCTGCGCGGCGGCTTGGCGGCA 3435
Qy 2103 GCTGAGTTCGGGGTGGGGTGGGGTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGG 2162
Db 3436 GCTGAGTTCGGGGTGGGGTGGGGTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGG 3495
Qy 2163 CTGCGAGGCGCGCGGAGATGAGAGGCTTGAACCCGCGGCGAGGGGCTGGGGGCTTAC 2222
Db 3496 CTG-CTGCGCGCTTGGCAATGAGAGGCTTGGCACCCGCGGCGAGGGCTGGAGAGGGTAC 3554
Qy 2223 TGGGTGCGCCAGCAGTGCAGCCCGCGGCGTGTGCTGCTCGATTTCTCATGAGGCT 2282
Db 3555 TGGGTGCGCCAGCAGTGCAGCCCGCGGCGTGTGCTGCTCGATTTCTCATGAGGCT 3610
Qy 2283 TTAGCAGCCTTCCCGCGGCGGAGGCTGGGAGCTTGAAGCCCGCATGCTGAGGCT---- 2338
Db 3611 TTAGCCTCTTCCCGCGGCGGAGGCTGGGAGCTTGAAGCCCGCATGAGGCTTGAAGCC 3670
Qy 2339 -CCCTTCATGAGGCTTCTGCGGCGGAGGCTTCCCGGAGAG-CACACACCCCTGCTCC 2396
Db 3671 CCCCCATGAGGCTCTGTGCTGCGGAGGCTTCCCGGAGAGTGCACACCCCGCTCC 3730
Qy 2397 ACAGGCGCCAGTCCCATGACACAGCAGAGGCTGAGAGTGGCGGCGCAGCGGCGAG 2456
Db 3731 AGGCGCGCCAGTCCCATGACACAGCAGAGGCTGAGAGTGGCGGCGCAGCGGCGAG 3789
Qy 2457 CTGGGAGGAGGTACCCCTGAGCCCTGAGGCGGAGTACATGGGAGGAGCCAGCTGGGC 2516
Db 3790 CTGGTGGGAGGTCCACCTGAGCCCGCGGAGGAGTCCATGGGAGGAGCCAGCTGGGC 3849
Qy 2517 TCCCTAGTCTGAGTGGAGACTTGGAGACCTTATGTCTAGCTCAGGAGTCTGTAATAC 2576
Db 3850 TCCCTAGTCTGAGTGGGAGGAGTGGAGAACCTTATGTCTAGCTCAGGAGTCTGTAATA 3906
Qy 2577 CAATCAGCACCTGTGTCTAGCTCAGGAGTCTGTAATGACACCAATCAGCTGTATCT 2636
Db 3907 -----CACCAATGACACCTGTATCT 3928
Qy 2637 ACCTACTGATGGGCGCTTGGAGAACCTTATGTCTAGCTCAGGAGTCTGTAATACAC 2696
Db 3929 ACCTACTGATGGGCGCTTGGAGAACCTTATGTCTAGCTCAGGAGTCTGTAATACAC 3988
Qy 2697 AATCGGCACTCTGTATCTAGCTCAGAGTGTGTAACACACCAATCAGCACCTGTGTCTA 2756
Db 3989 AATCAGCACTCTGTATCTAGCTCAGAGTGTGTAACACACCAATCAGCACCTGTGTCTA 4048

Qy 2757 GCTCAGGGTATGTGAATGACCAATGACAGTCTGTATCTGGCTACTTTCATGGG 2811
Db 4049 GCTCAGGGTGTGTAATGACCAATGACAGTCTGTATCTGGCTACTTGTGGG 4103
RESULT 14
US-60-487-610-19299
; Sequence 19299, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19299
; LENGTH: 70556
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(70556)
; OTHER INFORMATION: n = A,T,C or G, or Insertion/deletion polymorphism (see Tables
US-60-487-610-19299
Query Match 10.9%; Score 677.2; DB 7; Length 70556;
Best Local Similarity 67.1%; Pred. No. 5.4e-101;
Matches 1366; Conservative 6; Mismatches 254; Indels 410; Gaps 13;
Qy 1862 TGGCTTGGAGCTGAGAGGTGACAGCATGCGCGGAGTCTCAGACGCCCTGCTCTCG 1921
Db 3120 TGAATACATCTGCTGAGGAGTACAGCTGCTGAGTCTCAGAGCCC---TGGCTCTCG 31255
Qy 1922 GCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1981
Db 31256 GCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31315
Qy 1982 TGCATGTGTGAGAGCCCTTCTTGTGGCTGGCCAAAGCCGAGCCGCTTCACTTGTGA 2041
Db 31316 TGCATGTGTGAGAGCCCTTCTTGTGGCTGGCCAAAGCCGAGCCGCTTCACTTGTGA 31375
Qy 2042 GGAAGGTGTGAGAGGAGAGGCTCAAGCAGAGACCGGGGCTGGCGAGCGGCTTGGCGGC 2101
Db 31376 GGAAGGTGTGAGAGGAGAGGCTCAAGCAGAGACCGGGGCTGGCGAGCGGCTTGGCGGC 31429
Qy 2102 ACCTGAGTTCGCGGCTGGGCTGGGCTTGGCGGCGCCGACACTGCGAGCAGCGGCGCAGC 2161
Db 31430 ACCTGAGTTCGCGGCTGGGCTGGGCTTGGCGGCGCCGACACTGCGAGCAGCGGCGCAGC 31489
Qy 2162 CCTGCCAGGCGCCCGGAGTGAAGAGCTTGAAGACCCGCGGCGAGCGGCTTGGAGGTTGA 2221
Db 31490 CCTTCTGCGGCGGAGTGAAGAGCTTGAAGACCCGCGGCGAGCGGCTTGGAGGTTGA 31548
Qy 2222 CTGGGAGCCCGGAGTGCAGACCCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2281
Db 31549 GTGAGTCTCTCAGAGTGCAGACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 31604
Qy 2282 TTAGCAGCTTCCCGGCGGAGGCTGCGGAGCTTGCAGACCCGCGCAGTGCCTTGAAGCT 2338
Db 31605 TTAGCTGCTTCTGCGGCGGAGGCTGCGGAGCTTGCAGTGCCTGCTGCTGCTGCTGCTG 31664
Qy 2339 --CCCTTCATGAGGCTCTGTGCGGCGGAGGCTTCCCGGAGCAGACACCCCTGTCTCC 2396
Db 31665 ACCCCCTCATGAGGCTCTGTGCGGCGGAGGCTTCCCGGAGCAGACACCCCTGTCTCC 31724
Qy 2397 ACAGGCGCCAGTCCCATGACACAGCAGAGGCTGAGAGTGGGCGGAGCAGGCGCAGGCGGA 2456
Db 31725 ACCGACCCAGTCCCATGACACAGCAGAGGCTGAGAGTGGGCGGAGCAGGCGCATGGCA 31783


```
Db 55667 GTGTGGAGAGAGAGCCGAGCGGGAACAGGGCTGCGCGGGGCTTGACAGGCCAGCTG 55626
QY 2107 GAGTTCGGGGTGGCGCTGGGCTTGGCGGGGCCGCACTGGAGAGCGGGCCAGCCCTGC 2166
Db 55627 GAGTTCGGCGTGGCGCTGGGCTTGGC-GGCCCGGCACTGGAGAGCAACGGCCGCTGC 55685
QY 2167 CAGGCCCGGGCAATTGAGAGGCTTACACCCGGGGCGAGCGGCTGGAGAGGCTGTAGTGG 2226
Db 55686 C-GGCCCGGGCAATTGAGAGGCTTACACCCGGGGCGAGCGGCTGAGAGAGGCTGTAGTGG 55744
QY 2227 TGCCCCAGCAGTGGCAGCCCGCGGCTGTGCTGCTCGATTCTCAGTGGGCTTAGC 2286
Db 55745 TCCCCAGCAGTGGCAGCCCGCGGCTG---CGCTCGACTTCTCGCGGCTTAGC 55800
QY 2287 AGCCTTCGGGGGGGAGGGGCTGGGACCTGGACGCGGCCCATGCTGAGGCTCC----- 2341
Db 55801 TGCTTTCGGGGGGGAGGGCTGGGACCTGGACGCGGCCCATGCTGAGGCTCCACCCG 55860
QY 2342 CTCATGGGCTCTGTGCGGGCCGAGCCTCCCGCAGCAGCACACCCCTGCTCCAGC 2401
Db 55861 CTCGCTGGGCTCTGTGCGGGCCGAGGCTTCCGATAGGCGCGCCCTGCTGACGGC 55920
QY 2402 GCCCAGTCCATCGACCAAGCAGGGCTGAGAGTGGGGCGCACGGCACCGGACTGGC 2461
Db 55921 ACCCAGTCCATCGACCAAGCAGGGCTGAGAGTGGCGGCG-----TGAGAGTGGC 55972
QY 2462 AGCAGCTACCCCTGCGAGCCCTGGTGGGAATCCAGTGGGGTGAAGCCAGTGGGCTCTG 2521
Db 55973 AGGTAGCTCCGCTGCGAGCCCGGTGGGATCCAGTGGGTAAGCCAGTGGGCTCTG 56032
QY 2522 AGTCTGGTGGAGACTTGGAGAACCTTATGTCTAGCTCAGGGATCGTAAT----- 2572
Db 56033 AGTCTGGTGGGAGCTGGAGAACCTTACCTAGCTCAGGGATCGTAATACAAATC 56092
QY 2573 -----ACCAATCAGCACCCCTGTCTAGCTC 2600
Db 56093 AGCAGCTGTGTCTAGCTCAAGGTTGTAAACACACCAATCAGCACCCCTGTCTAGCTC 56152
QY 2601 AGGCTGTGTGATGACCAATCAGCACTGTATCTA-----GTACTGTGATGGGCC 2654
Db 56153 AGGCTTGTGATGACCAATCAGCACTGTATCTACTGTACTGTGTGTGGGCC 56212
QY 2655 TTGGAGAACCTTATGTCTAGCTCAGGATTTGTAATACAAATCGGCACTGTATCT 2714
Db 56213 TTGGAGAACCTTATGTCTAGCTCAGGATTTGTAATACAAATCGGCACTGTATCT 56272
QY 2715 AGCTCAAGTTTGTAAACACACCAATCAGCACCCCTGTGTAGCTCAGGGTATGTGAATG 2774
Db 56273 AGCTCAAGTTTGTAAACACACCAATCAGCACCCCTGTGTAGCTCAGGGTATGTGAATG 56332
QY 2775 CACCAATCAGAGTGTATCT 2796
Db 56333 CACCAATCAGACTGTGATCT 56354
```

Search completed: August 17, 2003, 01:39:35
Job time : 648 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2003, 13:22:16 ; Search time 7800 Seconds
(without alignments)
19428.001 Million cell updates/sec

Title: US-09-845-020A-5
Perfect score: 6235
Sequence: 1 gatcacttgagacagtagt.....tgcacatcccccacagcgct 6235

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estdb:*
2: em_esthm:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	711.8	11.4	3969 11	BC033022 Homo sapi
2	710.4	11.4	1066 13	BX423847 BX423847
3	707.8	11.4	1201 9	AL532400 AL532400
4	707.2	11.3	2067 11	AF318322 Homo sapi

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
BC033022	BC033022	Homo sapiens, clone IMAGE:4828044, mRNA.	3969 bp	mRNA	11near	HTC	04-MAR-2003						
BC033022.1	BC033022.1	GI:22766856											
Homo sapiens (human)													
Homo sapiens													
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.													
1 (bases 1 to 3969)													
Strausberg, R.													
Direct Submission													
Submitted (21-JUN-2002) National Institutes of Health, Mammalian													
Gene Collection (MGC), Cancer Genomics Office, National Cancer													
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,													
USA													
NIH-MGC Project URL: http://mgc.ncl.nih.gov													
Contact: MGC help desk													
Email: cgapbs-remail.nih.gov													
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.													
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shitaki													
Toshiyuki and Piero Carninci (RIKEN)													
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)													
DNA Sequencing by: Institute for Systems Biology													
http://www.systemsbiology.org													

ALIGNMENTS

C	5	697.8	11.2	1201	13	BX358085	BX358085
C <td>6</td> <td>692.8</td> <td>11.1</td> <td>1201</td> <td>13 <td>BX402423</td> <td>BX402423</td> </td>	6	692.8	11.1	1201	13 <td>BX402423</td> <td>BX402423</td>	BX402423	BX402423
C <td>7</td> <td>690</td> <td>11.1</td> <td>955</td> <td>13 <td>BX407975</td> <td>BX407975</td> </td>	7	690	11.1	955	13 <td>BX407975</td> <td>BX407975</td>	BX407975	BX407975
C <td>8</td> <td>687.8</td> <td>11.0</td> <td>1201</td> <td>13 <td>BX402424</td> <td>BX402424</td> </td>	8	687.8	11.0	1201	13 <td>BX402424</td> <td>BX402424</td>	BX402424	BX402424
C <td>9</td> <td>670.4</td> <td>10.8</td> <td>1201</td> <td>13 <td>BX358084</td> <td>BX358084</td> </td>	9	670.4	10.8	1201	13 <td>BX358084</td> <td>BX358084</td>	BX358084	BX358084
C <td>10</td> <td>600</td> <td>9.6</td> <td>879</td> <td>13 <td>BX346821</td> <td>BX346821</td> </td>	10	600	9.6	879	13 <td>BX346821</td> <td>BX346821</td>	BX346821	BX346821
C <td>11</td> <td>593.6</td> <td>9.5</td> <td>1914</td> <td>11</td> <td>AF289565</td> <td>AF289565</td>	11	593.6	9.5	1914	11	AF289565	AF289565
C <td>12</td> <td>486.8</td> <td>7.8</td> <td>1060</td> <td>29</td> <td>BX598357</td> <td>BX598357</td>	12	486.8	7.8	1060	29	BX598357	BX598357
C <td>13</td> <td>479.6</td> <td>7.7</td> <td>816</td> <td>10</td> <td>BF344606</td> <td>BF344606</td>	13	479.6	7.7	816	10	BF344606	BF344606
C <td>14</td> <td>478</td> <td>7.7</td> <td>631</td> <td>10</td> <td>AM859880</td> <td>AM859880</td>	14	478	7.7	631	10	AM859880	AM859880
C <td>15</td> <td>461.2</td> <td>7.4</td> <td>723</td> <td>13 <td>BX095590</td> <td>BX095590</td> </td>	15	461.2	7.4	723	13 <td>BX095590</td> <td>BX095590</td>	BX095590	BX095590
C <td>16</td> <td>460.8</td> <td>7.4</td> <td>806</td> <td>29</td> <td>BZ600353</td> <td>BZ600353</td>	16	460.8	7.4	806	29	BZ600353	BZ600353
C <td>17</td> <td>458.6</td> <td>7.4</td> <td>1201</td> <td>13 <td>BX343659</td> <td>BX343659</td> </td>	17	458.6	7.4	1201	13 <td>BX343659</td> <td>BX343659</td>	BX343659	BX343659
C <td>18</td> <td>457.2</td> <td>7.3</td> <td>943</td> <td>13 <td>BX349933</td> <td>BX349933</td> </td>	18	457.2	7.3	943	13 <td>BX349933</td> <td>BX349933</td>	BX349933	BX349933
C <td>19</td> <td>456</td> <td>7.3</td> <td>629</td> <td>28</td> <td>AQ041508</td> <td>AQ041508</td>	19	456	7.3	629	28	AQ041508	AQ041508
C <td>20</td> <td>452</td> <td>7.2</td> <td>604</td> <td>12</td> <td>BM008701</td> <td>BM008701</td>	20	452	7.2	604	12	BM008701	BM008701
C <td>21</td> <td>443</td> <td>7.1</td> <td>909</td> <td>10</td> <td>BM675935</td> <td>BM675935</td>	21	443	7.1	909	10	BM675935	BM675935
C <td>22</td> <td>439.4</td> <td>7.0</td> <td>567</td> <td>12</td> <td>BM706947</td> <td>BM706947</td>	22	439.4	7.0	567	12	BM706947	BM706947
C <td>23</td> <td>436.4</td> <td>7.0</td> <td>846</td> <td>14</td> <td>CD171662</td> <td>CD171662</td>	23	436.4	7.0	846	14	CD171662	CD171662
C <td>24</td> <td>433</td> <td>6.9</td> <td>763</td> <td>28</td> <td>AQ371716</td> <td>AQ371716</td>	24	433	6.9	763	28	AQ371716	AQ371716
C <td>25</td> <td>429</td> <td>6.9</td> <td>977</td> <td>14</td> <td>CD051193</td> <td>CD051193</td>	25	429	6.9	977	14	CD051193	CD051193
C <td>26</td> <td>428.6</td> <td>6.9</td> <td>948</td> <td>10</td> <td>BF795432</td> <td>BF795432</td>	26	428.6	6.9	948	10	BF795432	BF795432
C <td>27</td> <td>423</td> <td>6.8</td> <td>989</td> <td>13</td> <td>BQ691867</td> <td>BQ691867</td>	27	423	6.8	989	13	BQ691867	BQ691867
C <td>28</td> <td>421.6</td> <td>6.8</td> <td>678</td> <td>29</td> <td>AG064242</td> <td>AG064242</td>	28	421.6	6.8	678	29	AG064242	AG064242
C <td>29</td> <td>420.4</td> <td>6.7</td> <td>674</td> <td>29</td> <td>AG097408</td> <td>AG097408</td>	29	420.4	6.7	674	29	AG097408	AG097408
C <td>30</td> <td>417.6</td> <td>6.7</td> <td>743</td> <td>29</td> <td>AG088193</td> <td>AG088193</td>	30	417.6	6.7	743	29	AG088193	AG088193
C <td>31</td> <td>414.2</td> <td>6.6</td> <td>711</td> <td>14</td> <td>CB962291</td> <td>CB962291</td>	31	414.2	6.6	711	14	CB962291	CB962291
C <td>32</td> <td>411</td> <td>6.6</td> <td>900</td> <td>13</td> <td>BQ166056</td> <td>BQ166056</td>	32	411	6.6	900	13	BQ166056	BQ166056
C <td>33</td> <td>408.6</td> <td>6.6</td> <td>549</td> <td>28</td> <td>AQ411493</td> <td>AQ411493</td>	33	408.6	6.6	549	28	AQ411493	AQ411493
C <td>34</td> <td>405</td> <td>6.5</td> <td>1201</td> <td>9</td> <td>AL541354</td> <td>AL541354</td>	34	405	6.5	1201	9	AL541354	AL541354
C <td>35</td> <td>403</td> <td>6.5</td> <td>681</td> <td>29</td> <td>AG103225</td> <td>AG103225</td>	35	403	6.5	681	29	AG103225	AG103225
C <td>36</td> <td>402.6</td> <td>6.5</td> <td>647</td> <td>29</td> <td>AG177729</td> <td>AG177729</td>	36	402.6	6.5	647	29	AG177729	AG177729
C <td>37</td> <td>402.6</td> <td>6.5</td> <td>920</td> <td>13</td> <td>BQ552848</td> <td>BQ552848</td>	37	402.6	6.5	920	13	BQ552848	BQ552848
C <td>38</td> <td>401.4</td> <td>6.4</td> <td>955</td> <td>10</td> <td>BF795502</td> <td>BF795502</td>	38	401.4	6.4	955	10	BF795502	BF795502
C <td>39</td> <td>400</td> <td>6.4</td> <td>635</td> <td>12</td> <td>B1830860</td> <td>B1830860</td>	39	400	6.4	635	12	B1830860	B1830860
C <td>40</td> <td>398.4</td> <td>6.4</td> <td>516</td> <td>28</td> <td>AQ746239</td> <td>AQ746239</td>	40	398.4	6.4	516	28	AQ746239	AQ746239
C <td>41</td> <td>396.4</td> <td>6.4</td> <td>819</td> <td>10</td> <td>BF813114</td> <td>BF813114</td>	41	396.4	6.4	819	10	BF813114	BF813114
C <td>42</td> <td>396.2</td> <td>6.4</td> <td>907</td> <td>13</td> <td>BQ890552</td> <td>BQ890552</td>	42	396.2	6.4	907	13	BQ890552	BQ890552
C <td>43</td> <td>396.2</td> <td>6.4</td> <td>949</td> <td>13</td> <td>BQ716016</td> <td>BQ716016</td>	43	396.2	6.4	949	13	BQ716016	BQ716016
C <td>44</td> <td>396</td> <td>6.4</td> <td>507</td> <td>28</td> <td>AQ415290</td> <td>AQ415290</td>	44	396	6.4	507	28	AQ415290	AQ415290
C <td>45</td> <td>396</td> <td>6.4</td> <td>932</td> <td>13</td> <td>BX348654</td> <td>BX348654</td>	45	396	6.4	932	13	BX348654	BX348654

contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRK Plate: 34 Row: c Column: 24
This clone has the following problem: retained Intron.

FEATURES
source

1.3969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4828044"
/issue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1047 a 928 c 902 g 1092 t
ORIGIN

Query Match 11.4%; Score 711.8; DB 11; Length 3969;
Best Local Similarity 87.9%; Pred. No. 8.2e-07;
Matches 824; Conservative 0; Mismatches 102; Indels 11; Gaps 4;

QY 1869 AGCTGAGAGGTACAGCATGCCGCAAGTCTCTACAGCCCTCGTTCGCGCCCTC 1328
DB 411 AGACCCAGAGGTACAGCATGCTGCGACGCTCTCAGAGCCCTTCCTGCTCCGCACTC 470
QY 1929 CTCTGCTGGGCTCCCACTTGGTGGCACTTGGAGAGCCCTTACGCCACCGTGACTG 1988
DB 471 CCTGCTGGGCTCCCGCTTTGGTGGCACTTGGAGAGCCCTTACGCCACCGTGACTG 530
QY 1989 TGGGAGCCCTTCTGCTGGTGGCAAGGCCGAGCCGCTCTCTCACTTGCAGGAGGT 2048
DB 531 TGGGAGCCCTTCTGCTGGTGGCAAGGCCGAGCCGCTCTCTCACTTGCAGGAGGT 590
QY 2049 GTGGAGGAGAGGCTCAAGCAGAACCGGGGCTGCGCAGCCCTTGGGGCCAGCTGGA 2108
DB 591 GTGGAGGAGAGGCTCAAGCAGAACCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 650
QY 2109 GTTCCGGGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTT 2168
DB 651 GTTCCGGGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTT 709
QY 2169 GGGCCCGGGCAATGAGAGGCTTACACCGGGGCAAGCGGCTGCGAGGGGTACTGGGTG 2228
DB 710 GGGCCCGGGCAATGAGAGGCTTACACCGGGGCAAGCGGCTGCGAGGGGTACTGGGTG 769
QY 2229 CCGCAGAGTCCAGCCCGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2288
DB 770 CCGCAGAGTCCAGCCCGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 2289 CCGTCCCGGGGCAAGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 2343
DB 826 CCGTCCCGGGGCAAGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 885
QY 2344 CCGTGGGCTCTGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 2403
DB 886 CCGTGGGCTCTGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 945
QY 2404 CCGATGCCATTCGACACGCAAGGGGTGAGGAATGCGAGGCCCGGCA-CAGGACTGGCAG 2463
DB 946 CCGATGCCATTCGACACGCAAGGGGTGAGGAATGCGAGGCCCGGCA-CAGGACTGGCAG 1004
QY 2464 GGAGCTACCCCTCCAGCCCTGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 2523
DB 1005 GGAGCTACCCCTCCAGCCCTGTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 1064
QY 2524 TGTGGTGGAGACTTGGAGAACCTTGTATGCTAGCTAGAGGATGTAATACCAATTCAG 2583
DB 1065 TGTGGTGGAGACTTGGAGAACCTTGTATGCTAGCTAGAGGATGTAATACCAATTCAG 1124

QY 2584 CACCGTGTCTAGCTCAGGCTCTGGAATGACCAATCCACACTCTGTATAGTACT 2643
DB 1125 CACCGTGTCTAGCTCAGGCTCTGGAATGACCAATCCACACTCTGTATAGTACT 1184
QY 2644 CTGATGGGCGCTTGGAGAACCTTGTATGCTAGCTCAGGAGTGTAAATACCAATCCAGC 2703
DB 1185 CTGATGGGCGCTTGGAGAACCTTGTATGCTAGCTCAGGAGTGTAAATACCAATCCAGC 1244
QY 2704 ACTGTGTATCTAGCTCAGGCTTGTAAACACCAATCAGACCCGTGTCTAGCTCAGG 2763
DB 1245 ACTGTGTATCTAGCTCAGGCTTGTAAATACCAATCAGACCCGTGTCTAGCTCAGG 1304
QY 2764 GTATGTGAATGACCAATTCAGACTCTGTATCTAGCTCAGG 2800
DB 1305 GTTGTGAATGACCAATTCAGACTCTGTATCTAGCTCAGG 1341

RESULT 2

LOCUS BX423847 1066 bp mRNA linear EST 13-MAY-2003
DEFINITION BX423847 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
ACCESSION CS0DM014YB01 5-PRIME, mRNA sequence.
VERSION BX423847
KEYWORDS BX423847.1 GI:30659589
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE 1 (bases 1 to 1066)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : filang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Faraday Avenue Genoscope sequence ID : CS1AM0042D010P1.

FEATURES
source

1.1066
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM014YB01"
/issue_type="FETAL LIVER"
/dev_stage="Fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: Liver; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 200 a 320 c 314 g 223 t 9 others
ORIGIN

Query Match 11.4%; Score 710.4; DB 13; Length 1066;
Best Local Similarity 85.3%; Pred. No. 2.4e-06;
Matches 834; Conservative 6; Mismatches 127; Indels 11; Gaps 4;

QY 1859 GGAATTCCTGAGGCTGAGAGGTACAGCATGCGCGAGTCTCAGAGCCCTGCTGGCTC 1918
DB 50 GGAATTCCTGAGGCTGAGAGGTACAGCATGCGCGAGTCTCAGAGCCCTGCTGGCTC 109
QY 1919 TCGGCGCTCTCTGCTGGGCTCCCACTTGGTGGCACTTGGAGAGCCCTTCAAGCCAC 1978
DB 110 TCGGCGCTCTCTGCTGGGCTCCCGCTTGGTGGCACTTGGAGAGCCCTTCAAGCCAC 169
QY 1979 CGGTGACACTGTGGAGACCCCTTCTGGGCTGGCGCAAGGCCGCGCTCCCTCAGCTT 2038
DB 170 CACTGCACTGTGGAGACCCCTTCTGGGCTGGCGCAAGGCCGCGCTCCCTCAGCTT 229

QY	2039	GCAGGAGAGTGTGGAGGAGAGGCTCAAGCAGAAACCGGGGCTGCGACGGCGCTTTGGCG	2098
Db	230	GCAGGAGAGTGTGGAGGAGAGACACGAGCGGGAACCGGGGCTGTCTGCACCTTGGCG	289
QY	2099	GCAGGCTGGAATCCGGGTGGGGCGTGGGGCTTGGCGGGCCCGGCATTCGGAGCAGGGGCC	2155
Db	290	GCATCTTGAGTCCGGGTGGGGCGTGGGGCTTGGTGGGCCCGGCATTCAGACAGCACCC	349
QY	2159	AGCCCTGCGAGAGCCCGGGGCATGAGAGGCTTTCAGCACCAGGAGCAGCGGCTGCGGAGGT	2218
Db	350	AGCCCTG - CTGGCCCCCGGCATGGGGGACTTACACTGGGGCAGTGGCTGCGGAGGCT	408
QY	2219	GTACTGGGTGCCCCAGCAGTGTCCAGCCCGCGCGCTTGCTCGCTGATTTTCATCG	2278
Db	409	GTACTGATTCGCCAGCAGTGTGGCCGACCGGGCTG --- CGCTCATTTCTTCGCTGG	464
QY	2279	GCTTTAGAGGCTTCCCGGGGGCGGAGGCTGCGGGACCTGCGACGCCCGCATCGTAGGCT	2338
Db	465	GCTTTGGTGGCTTCCCGAGGGGGAGGAGGCTGCGGAGCTGCGAGCAGCACCCCTGCG	524
QY	2339	-----CCCTTCATGGGCTCCTGTGGGGCCGAGGCTCCCGAGAGCAGCACCCCTGCG	2393
Db	525	CCTACCCGATCCATGGGGCTCCTGTGGGGCTGAGAGGCTCCCGAGCAGCACCCCTGCG	584
QY	2394	TCCACAGAGGCCAGTCCCATGAGCAGCAGCAAGGGCTGGAATGGGGGGCGACAGGACCG	2455
Db	585	TCCACAGGGGCCCATGCTCCATGAGCAGCAGCAAGGGCTGGAATGGAGCGGCCCGCA -CA	643
QY	2454	GGACTGGCAGGAGCTACCCCTGCGAGCCCTGGTGGGAACTCATCGGGTGAAGCAGCTG	2513
Db	644	GGACTGGCAGGAGCTACCCCTGCGAGCCCGCATGGAGATGACATAGTGAAGCCAGCTG	703
QY	2514	GGCTTCGAGTCTGTGGAGAGCTGGAGAACTTTATGTCTAGCTCAGGAGATCGTAATA	2573
Db	704	GGCTTCGAGTCTGTGGAGAGCTGGAGAACTTTATGTCTAGCTCAGGAGATCGTAATA	763
QY	2574	CACCAATCAGACCCCTGTGTAGCTCAGGGGTGTGTGATGCACCAATCCACCTCTGTA	2633
Db	764	CACCAATCAGACCCCTGTGTAGCTCAGGGGTGTGTGATGCACCAATCCACCTCTGTA	823
QY	2634	TCTAGCTACTGTATGGGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGATGTAATAC	2693
Db	824	TCTAGCTACTGTGTGGAGGGGCTGGAGAGCTTTATGTCTAGCTCAGGAGATGTAATAC	883
QY	2694	ACCAATCGGAGCTGTGTATAGCTCAAGTTTGAACACACCAATCAGGAGACCTGTGT	2753
Db	884	ACCAATCGGAGCTGTGTATAGCTCAAGTTTGAATTAACCAATCGGAGCTGTGTAT	943
QY	2754	CTAGCTCAGGGTATGTGAATGCACCAATCAGCAGCTGTATCTGGCTACTTTCAGGGCA	2813
Db	944	CTAGCTCAGGGTATGTGAATGCACCAATCAGCAGCTGTGTATTCAGGGTGTGTAG	1003
QY	2814	TCCGCTGGAGAGACAC 2831	
Db	1004	TCCGCTGGAGAGACCTC 1021	
RESULT 3	AL532400	1201 bp	EST 23-MAY-2003
LOCUS	AL532400	Homo sapiens FETAL LIVER Homo sapiens cDNA clone	
DEFINITION	AL532400	5-PRIME, mRNA sequence.	
ACCESSION	AL532400		
VERSION	AL532400.2	GI:31070232	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Eumalpia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	L1, W.B., Gruber, C., Jessee, J. and Polayes, D.		
	Full-length cDNA libraries and normalization		

Unpublished
On Feb 13, 2001 this sequence version replaced gi:12795893.
JOURNAL
COMMENT

Unpublished
On Feb 13, 2001 this sequence version replaced gi:12795893.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@life tech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODM014CA010P1.

FEATURES
Source
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM014YB01"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_1lib="Homo sapiens FETAL LIVER"
/note="Organ : Liver; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT 229 a 345 c 350 g 254 t 23 others

ORIGIN

Query Match 11.4%; Score 707.8; DB 9; Length 1201;
Best Local Similarity 87.0%; Pred. No. 2.5e-06;
Matches 824; Conservative 2; Mismatches 109; Indels 12; Gaps 4;

1859 GGATGCGCTTGAAGCCCGAGAGGTGACAGATACCGGCATCTCTACAGCCCTCGTTGCGTC 191B
|||
49 GGAATTCCTCCGGAGATCGAGAGGTGACAGCATCTGCACTCTCTCAAGCCCTTGTGCTC 108
1919 TCGGCGCCTCCCTCTGCTGCGGCTCCCACTTGCGTGCGCACTTGAGAGGCCCTTCAAGCCAC 197B
109 TCGGCACTCTCCCTGCGCTGGGCTCCCGCTTTGTGTGCACTTGAAGAGGCCCTTCAAGTCCC 168
1979 CGCTGCACTGTGGGAGCCCTTTTGTGGGCTGGCCAGGCCAGAGCCGCTCCCTACGTT 203B
|||
169 CACTGCACTGTGGGAGCCCTTTTGTGGGCTGGCCAGGCCAGAGCCCTCCCTACGTT 228
2039 GCAGGAGAGTGTGGAGGAGAGGCGTCAAGAGAAACGGGGCTGGCGAGGGCGGCTTGGCG 209B
229 GCAGGAGAGTGTGGAGGAGAGACACAGACAGCCGGAAACGGGGCTGTGTCTGCACTTGGCG 288
2099 GCCAGCTGAGATCCGGGTGGGCGTGGGCTTGGCGGGGCCCGCACTCGAGAGCAGGGGCC 215B
289 GCCATCTGAGATTCGGGGTGGGCGTGGGCTTGTGGGGCCCGCACTCAGACAGCAGCACCC 348
2159 AGCCCTGCGCAGGCCCCGGGGCAATGAGAGGCTTACAGCCCGGGCCAGCGGGCTGGGAGGTT 221B
349 AGCCCTGCG -TGCCCCGGGGCAATGAGGAGCTTACCACTTGGGCCAAGTGCGCGGAGGTT 406
2219 GTACTGGGTGGCCAGAGAGTGCACAGCCCGCGGGCTGTGCTGCTGATTTTCACATGG 227B
407 GTACTGAGTCCCCCAGCAGTGTCTGGCCACCGGGCGTG -CGCTGATTTTCTGCGTGG 462
2279 GCGTTAGCAGCTTCCCGGGGGGAGGAGGCTTGGGAGCTGAGCCCGGCATGAGGCT 233B
463 GCGTTGAGCTGCTTCCACAGGGGAGGAGGCTTGGGAGCTGAGCCCGGCATGAGGCT 522
2339 CCCC-----TCGATGAGGCTCTGTGTGGGCGCCGAGACCTCCCGAGAGAGACACCCCGCTGC 2393B
|||
523 CTCATCCCAATCTCAATGAGGCTCTGTGTGGGCGCTGAGACCTCCCGAGAGAGACACCCCGCTGC 582
2394 TCCACAGCGCCCATCTCCATGACAGCAGCAAGAGGCTAGAAGTGCAGGCGCAGCGACGCAAG 2453B
583 TTCAGAGGCGCCCATCTCCATGACAGCAGCAAGAGGCTAGAAGATGCGAGCGCCCGGCA-CA 641
2454 GGACTGCGAGCAGCTACCCCTGTGAGCGCCCTGTGGTGGGAATTCACCTTGGGTGAAGCGAGCTG 2513B

Db 642 GGACTGGCAGGAGCTCCACCTGCAGCCGCCAGTGCAGATCAGTACGTAGTGAAGCCAGCTG 701

QY 2514 GGCTCTGAGTGTGGTGGAGACTTGGAGAACCTTTATGTCTAGCTCAGGAGCTGTAATA 2573

Db 702 GGCTCTGAGTGTGGTGGAGAGTCTTTATCTAGCTCAGGAGATGTGAATA 761

QY 2574 CACCAATCAGCAACCTGTGTCTAGCTCAGGAGTCTGGAATGCACCAATCAGCACTGTGA 2633

Db 762 CACCAATCAGCAACCTGTGTCTAGCTCAGGAGTCTGGAATGCACCAATCAGCACTGTGA 821

QY 2634 TCTAGCTACTGTGATGGGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATAAC 2693

Db 822 TCTAGCTACTGTGATGGGCTTGGAGAGTCTTTATGTCTAGCTCAGGAGTGTGAATAAC 881

QY 2694 ACCAATCGGCACTCTATCTAGCTCAGGAGTGTGAATAACACCAATCAGCACTGTGT 2753

Db 882 ACCAATCAGCACTCTGTCTAGCTCAGGAGTGTGAATAACACCAATCAGCACTGTGTAT 941

QY 2754 CTAGCTCAGGAGTGTGATGGAGCAGCAATCAGCACTGTGTATCTGAGCT 2800

Db 942 CTAGCTCAGGAGTGTGTGAACACCAATCAGCACTGTGTATCTGAGCT 988

RESULT 4
AF318322 2067 bp mRNA linear HTC 01-JAN-2002

LOCUS AF318322
DEFINITION Homo sapiens PPI1662 mRNA, complete cds.
ACCESSION AF318322
VERSION AF318322.1 GI:18027735
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2067)
Jiang,H.Q., Zhou,X.M., Zhang,P.P., Huang,Y., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
TITLE Novel human cDNA clones with function of inhibiting cancer cell growth

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2067)
AUTHORS Jiang,H.Q., Zhou,X.M., Zhang,P.P., Huang,Y., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai 200032, P. R. China

FEATURES
source location/Qualifiers
1..2067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2067
/gene="PPI1662"
1093..1701
/gene="PPI1662"
/codon_start=1
/product="unknown"
/protein_id="AAL55829.1"
/db_xref="GI:18027736"
/translation="MGIGPALGAAGRPRAVRADLAPAGACRGTCTSPSSAGPA
LCILSRMALAAPPRGRARDLPQAVPEPTPRSGSCGARSPTSTPCSRAPSDHPR
AECEERTARDQAPPAAPVAPVADLPGEASVAPSDGVESPVLLRDCKHTNHQHPVSS
REVSAVPDLYLAALVGPRTFPISSGIVNPIGIVYLAQGL"

BASE COUNT 406 a 603 c 590 g 468 t

ORIGIN

Query Match 11.3%; Score 707.2; DB 11; Length 2067;
Best Local Similarity 89.1%; Pred.No.1.7e-06;
Matches 845; Conservative 0; Mismatches 88; Indels 15; Gaps 7;

QY 1873 TGAGAGGTGACAGCATCGCGGAGTCTCAGAGCCCTGCTGGCTGGCGCTCCTCT 1932

||||| ||||||| || ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 841 TGAGGGGTGACAGACGTCGTGGCAGTCTCAGTCCAGACGCCCTCGCTGCTCGGCGCTCCTCT 900

QY 1933 GGCTGGGCTCCCACTTCGCTGGCAGCTTGAAGAGCCCTTACAGCCACCGCTGCAGTCT -GG 1991

Db 901 GGCTGGGCTCCCACTTCGCTGGGCGCACTTGAAGAGCCCTTACAGCCACCGCTGCAGTGGG 960

QY 1992 GAGCCCTTTCGCTGGGCTGGGCGCAAGGCGCAGAGCGGCTCCCTGCTGCTTGAAGGAGGTG 2051

Db 961 GAGCCCTTTCGCTGGGCTGGGCGCAAGGCGCAGAGCGGCTCCCTGCTGCTTGAAGGAGGTG 1020

QY 2052 GAGGAGG -AGGCTCAGCAGAGAACCGGGGCTGCGCAGCGGCTTGCAGGCGCAGCTGAG 2109

Db 1021 GATCGAGAGAGGCGCGAAGCGGGGACCGGGGCTGCGCAGCGGCTTGCAGGCGCTGCGGCGCAGCTGAG 1080

QY 2110 TTCCGGGCTGGGCTGGGCTTGGCGGCGCCGCACTGAGAGCAGCGGCGCAGCCCTGCCAG 2169

Db 1081 TTCCGGGCTGGGCTGGGCTTGGCGGCGCCGCACTGAGAGCAGCGGCGGCGCTTGC -G 1139

QY 2170 GCGCGGGGCAATAGAGGGCTTACACCGCGGCGCAGC -GGCTGCGGAGGGGTACTGGGTG 2228

Db 1140 GCGCGGGGCAATAGAGGGCTTGGCACCGCGGCGCAGCGGCTGCGAGAGGGGTACTGGGTG 1199

QY 2229 CCCAGCAGTGCAGCGCCGCGCGCTGCTGCTGATTTCTCACTGGGCTTGAAG 2288

Db 1200 CCCAGCAGTGCAGCGCGCGCCACACTGCGCTG -GCTGATTTCTGCTGGGCTTGAAGT 1255

QY 2289 CTTTCCCGGGGCGAGGCTCGGAGCTGCAGCGCGCCACTGCTGAGACCTGCC -CT 2343

Db 1256 CTTTCCCGGGGCGAGGCTCGGAGCTGCAGCGCGCCGCTGAGACCTGCCACCGCTG 1315

QY 2344 CCATGGGCTCCTGTGGGCGCGGCGCTCCCGCAGAGCAACCGCCCTGCTCCAGCGC 2403

Db 1316 CCATGGGCTCCTGTGGGCGCGGCGCTCCCGCAGAGTACACCGCTGCTCCAGCGC 1375

QY 2404 CCAGTCCCATTCAGCACAGCAGGCTGAGAGTGGCGGCGCAGCGGCGGAGCTGGCAG 2463

Db 1376 CCAGTCCCATTCAGCACAGCAGGCTGAGAGTGGCGGCGCAGCGGCGGAGCTGGCAG 1434

QY 2464 GCAGTACCCCTGCAGCCCTGCTGGGAGATCCACTGGGTGAAGCAGCTGGGCTGCTGAG 2523

Db 1435 GCAGTACCCCTGCAGCCCGGCTGGGAGATCCACTGGGTGAAGCAGCTGGGCTGCTGAG 1494

QY 2524 TCTGGTGGAGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATAACCAATCAG 2583

Db 1495 TCTGGTGGAGCTTGGAGAGTCTTTATGTCTAGCTCAGGAGTGTGAATAACCAATCAG 1554

QY 2584 CACCTGTGTCTAGCTCAGGAGTGTGAATAACCAATCAGCTGATCTAGTACT 2643

Db 1555 CACCTGTGTGTAGCTCAGGAGTGTGAAGTGTGAGTGCACAGTGCACCTGTATCTAGTGTCT 1614

QY 2644 CTGATGGGCGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATAACCAATCAG 2703

Db 1615 CTGATGGGCGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTGTGAATAACCAATCAG 1674

QY 2704 ACTCTGTATCTAGCTCAGGAGTGTGAATAACCAATCAGACCTGTGTCTAGTCTAG 2763

Db 1675 ACTGTATCTAGCTCAGGAGTGTGAATAACCAATCAGACCTGTGTCTAGTCTAG 1734

QY 2764 GATGTGAATGCACCAATGCAGACTGTGTATGGCTAGTACTTCAATGAGG 2811

Db 1735 GTTTGTGAGTGCACCAATCAACACTGTGTATCTAGTCTGTGTGTGG 1782

RESULT 5
BX358085/c 1201 bp mRNA linear EST 05-MAY-2003

LOCUS BX358085
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION BX358085
VERSION BX358085.1 GI:30370184
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Db	121	GCCTCCTCGCCTTGGGCTGCTCTACTTTTGGCGGCACTTGAAGACCCTTCAGCCACCGCTG	180
QY	1984	CACGTGGGAGGCCCTTCTGTGGGCTTGGCCAAAGCCGAGCCGCGCTCCCTCAGCTTGCAG	2043
Db	181	CAGTGTGGGAGGCCCTTCTGTGGGCTTGGCCAAAGCCGAGCCGCGCTCTCTCAGCTTGCAGG	240
QY	2044	GAGGTGTGGAGGAGAGGCTCAAGAGAACCGGGGGCTTGCACAGCGGCTTGCGGGCGAG	2103
Db	241	GAGGTGTGGAGGAGAGGCGCGAGGAGAACCGGGGGCTGTGTGCGCGCTTTCGGCGAG	300
QY	2104	CTGAGATTCCGGGTGGGGCTGGGCTTGGCGGGGCCCCGCACTCGAGAGCGGGCCAGGCC	2163
Db	301	CTGAGATTGCGGGGTGGGGCTGGGCTTGGCGGGGCCCCGCACTCGAGAGCCGCGCGAGCC	360
QY	2164	TGCCAGGCCCGGGCAATGAGAGGCTTAGACACCCGGGCGAGCGGCTGCGGAGGTGTACT	2223
Db	361	TGCTGGCGGCGCTGGCAATGAGAGGACTTAGACACCCGGGCGAGAGCTGCGGAGGGGTGCT	420
QY	2224	GGGTGCCCCAGAGATGAGCAGCCCGCGGGCGCTGTCTGCTGCTCGATTTTCTACTGGGCTT	2283
Db	421	GGGTGCCCCAGAGATGAGCAGCCCGCGGGCGCTGTCTGCTGCTCGATTTTCTACTGGGCTT	476
QY	2284	AGCAGCCTTCCCGCGGGGCGAGGGCTCGGGAGCTTGCAGCGCGCCGATGCTGAGCGCTCC	2341
Db	477	AGCTGCTTCCCGGGGGGCGAGGGCTCGGGAGCTTGCAGCGCGCCGATGCTGAGCGCTCC	536
QY	2342	---CTCCATGGGCTCTCTGTGGGGCCGAGCCTCCCCGAGAGCACACACCCCTGCTCCAC	2398
Db	537	CCACTGCAAGGGCTCTCTGTGGGGCCGAGCCTCCCTCGAGAGCACACACCCCTGCTCCAC	596
QY	2399	AGCGGCCAGTCCCATCGACCAAGGAGGCTGAGAGTGGGGGCGACAGCGGAGCT	2458
Db	597	GGCGGCCAGTCCCATCGACCAAGGAGGCTGAGAGTGGGGGCGACAGCGGAGCT	655
QY	2459	GCGAGGAGCAGTACCCCTCGAGCCCTGTGCGGAAATCCAGTGGGTGAAGCCAGCTGGGCT	2518
Db	656	GCTAGGAGCTCAGTACCTCGAGCCCGGGGTGCGGAAATCCAGTGGGTGAAGCCAGCGGCT	715
QY	2519	CTGAGTCTGTGTGAGAGCTTGGAGAACCTTATGTCTAGCTCAGGAGATCGTAAATACACA	2578
Db	716	CTGATGTGTGTGAGGAGCTGAGAGAGTCTTATGTCTAGCTCAGGAGATCGTAAATACACA	775
QY	2579	ATCAGACACCGTGTCTAGAGCTCAGGAGTGTGGAATGACCAATCCACACTCTGTATCTAG	2638
Db	776	ATTCAGACACCTGTGTCTAGAGCTCAGGAGTGTGGAATGACCAATCCACACTCTGTATCTAG	833
QY	2639	CTACTCTATGGGGCTTGGAGAACCTTATGTCTAGCTCAGGAGTGTGGAATGACCAAA	2698
Db	834	CTGCTCTGTGTGGGCTTGGAGAACCTTATGTCTAGCTCAGGAGTGTGGAATGACCAAA	893
QY	2699	TGCGCAGCTCTATCTAGCTCAGGAGTGTGGAACACACCAATTCAGACCCGTGTCTAGC	2758
Db	894	TGCGCAGCTCTATCTAGCTCAGGAGTGTGGAACACACCAATTCAGACCCGTGTCTAGC	953
QY	2759	TCAGGAGTGTGGAATGACCAATTCAGACCTGTCTAT	2794
Db	954	TCAGGAGTGTGGAATGACCAATTCAGACCTGTCTAT	989
RESULT 7			
BX407975			
LOCUS	BX407975	955 bp	mrna
DEFINITION	BX407975	Homo sapiens FETAL LIVER Homo sapiens cDNA clone	EST 15-MAY-2003
ACCESSION	CSDDM014YB01.5-PRIME	mrna sequence.	
KEYWORDS	BX407975		
SOURCE	BX407975.1	GI:30768714	
ORGANISM	Homo sapiens (human)		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (bases 1 to 955)			

AUTHORS	TITLE	JOURNAL	COMMENT
L.H.B., Gruber,C., Jesse,J. and Polayes,D.	Full-length cDNA libraries and normalization	Unpublished	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0A8015E02QPL
FEATURES	source	location/Qualifiers	
	1..955	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODM014YB01"	
		/tissue_type="FETAL LIVER"	
		/dev_stage="fetal"	
		/clone_1lib="Homo sapiens FETAL LIVER"	
		/note="Organ: Liver; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."	
BASE COUNT	176 a	293 c	291 g 191 t 4 others
ORIGIN			
Query Match	11.1%; Score 690;	DB 13:	Length 955;
Best Local Similarity	87.7%;	Pred. No. 6.4e-06;	
Matches	798;	conservative	3; Mismatches 98; Indels 11; Gaps 4
OY	1859	GGATGCTTGAGCCTGAGAGGTGACAGCATGCCGCAGTCCTCAAGCCCTGTTGCGTC	1918
Dd	52	GGAATTCCCCGGATGAGAGGTGACAGCATGCTGCGACATGCTCAAGCCCTGTTGCGTC	1111
OY	1919	TCCGGGCTCTCTCTGCTGCTGGGCTCCCACTTGGGTGGCACTTGAGAGCCCTTCAGCCCA	1977
Dd	112	TCCGGACTCTCCCTGCTGCTGGGCTCCCGCTTGGTGGCACTTGAGAGCCCTTCAGCTCCC	171
OY	1979	CGCTGCATGTGGGAGCCCTTCTGGGCTGGCCAAAGCCAGAGCCGGCTCCCTCAGCTT	2038
Dd	172	CACCTGCATGTGGGAGCCCTTCTGGGCTGGCCAAAGCCGGAGCCCACTCCTCAGCTT	231
OY	2039	GCAGGGAGTGTGAGGAGAGGCTCAAGACAAGACCAGGCTCGCACAGCGCTTGGCG	2098
Dd	232	GCAGGGAGTGTGAGGAGAGGAGACAAGACGCGGGAACCGGGCTGTGCTCGACTTGGGG	291
OY	2099	GCCAGCTGGAATTCGGGCTGGGCTGGGCTTGGCGGGCCCGGCACTCGAGACAGCGGCC	2158
Dd	292	GCCATCTGGAATTCGGGCTGGGCTGGGCTTGGGCGGCCCGGCACTCGAGACAGCAACC	351
OY	2159	AGCCCTGCGAGAGCCCGGGGCAATGAGAGGCTTAGACACCGGGGCGACGCGGTCGGAGGGT	2218
Dd	352	AGCCCTG-CITGGCCCCGGGCAATGGGGACTTAGACCTGGGCGCAAGTGGCTGGAGGGT	410
OY	2219	GTACTGGGTGCCCCAGCAGTGCCAGACCCCGCGGGCTGTGCTGCTCATTTTCATCTGG	2278
Dd	411	GTACTGGATGCCCCAGAGAGTGTGGCCACCGGGGCTG-----CGCTGATTTTCGCTGG	466
OY	2279	GCTTAGCAGAGCTTCCCGCGGGGCGAGGGCTTGGGAGCTTCAGAGCCCGGCATGCTGAGAGCT	2338
Dd	467	GCTTAGCTGCTCTCCACAGGGGGGAGGGCTTGGGAGCTTCAGAGCCCGGCATGCTGAGAGCT	526
OY	2339	-----CCCTTCATAGGGCTCTGTGGGGCCCGAGACCTTCCCGAGAGACCAACCCCTGC	2393
Dd	527	CTTACCCAATTCATGAGGGCTCTGTGGGGCCCTGAGACCTTCCCGAGAGACCAACCCCTGC	586
OY	2394	TCCACAGCGGCATTCCTCATGACACGAGAAGGGCTAGAGATGCGGGGCGACAGGCAACG	2453
Dd	587	TCCACAGCGGCATTCCTCATGACACGAGAAGGGCTAGAGATGCGGGGCGACAGGCAACG	645
OY	2454	GGAATGGCAGGACCTACCCCTGCAGCCCTGTGGTGGCAATTCATCTGAGGTGAAGCAACTG	2513

Dd		646	GGATCGGAGGACACTCCACCCTGCAGACGCCCACTGCAGGATCCTACGTAGTGAACCCAGCTG	705
Oy		2514	GGCTCCTGAGTCTGGTGAGAGACTTGGAGAACCTTTATGTCTAGCTCAGGATTCGTAAATA	2573
Dd		706	GGCTCCTGAGTCTGGTGAGGAGACGTGGAGAAGTCTTATATCTAGCTCAGGAGTTGTAATA	765
Oy		2574	CACCAATCAGACACCTGTGTCTAGCTACAGGGTCTGTGATCATGCACCAATCCACTCTGTA	2633
Dd		766	CACCAATCAGACACCTGTGTCTAGCTACAGGGTCTGTGATCATGCACCAATCCACTCTGTA	825
Oy		2634	TCTGAGCTACTGATGGGGGCGCTTGGAGAACCTTTATGTCTAGCTCAGGAGTTGTAATAAC	2693
Dd		826	TCTGAGCTCTCTGTGTGAGGGCGGTGGAGGCTTTATGTCTAGCTCAGGAGTTGTAATAAC	885
Oy		2694	ACCAATCGGACACTGTATCTAGCTCAGGAGTTTGTAAACACACCAATCAGCACCTCTGT	2753
Dd		886	ACCAATCAGCACTGTGTCTAGCTCAGGAGTTTGTAAATACACCAATCGGACACTCTGTAT	945
Oy		2754	CTAGCTCAGG	2763
Dd		946	CTAGCTCAGG	955
RESULT 8				
BK402424/c				
LOCUS				
DEFINITION	BK402424	1201 bp	mRNA	linear EST 13-MAY-2003
ACCESSION	BK402424			
VERSION	BK402424			
KEYWORDS	BK402424.1	GI:30618809		
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 1201)			
JOURNAL	L.I.W.B., Gruber,C., Jeessee,J. and Polayes,D.			
COMMENT	Full-length cDNA libraries and normalization unpublished Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : filiang@life tech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Faraday Avenue genoscope sequence ID : CSIAI009ZB08Qp1. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI033YB14" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cdna was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cdna was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
BASE COUNT	268 a	303 c	332 g	234 t
ORIGIN				
Query Match				
Best Local Similarity	89.2%	Score 687.8:	DB 13:	Length 1201:
Matches	805:	Conservative	4:	Mismatches 78:
			Indels	15:
			Gaps	6:
Oy	1915	GCTCTGGGCGCTCTCTGCTGTGGGCTCCACTTCGATGCGACTTGAAGAGCCCTTCAGC	1974	
Dd	975	GCTCTCCCGCGCTCTGCTGTGGGCTCTTATTGCGGACCTTGAAGAGCCCTTCAGC	916	
Oy	1975	CCACCGCTGACACTGTGGAGCCCTTTTGGGCTGGCCAAAGCCAGAGCCGGCTCCCTCA	2034	

Db	915	CCACCGCTGCAC	CTGTGGGAGGCCCTTTCTGTGGGCTTGCCCAAGGCGCGAGACCTCTCTCTTCA	856
QY	2035	GCTTGCAGGAG	GTGTGGAGGAGAGGCTCAGCAGGAACCGGGGCTTGCCAGCGGCTT	2094
Db	855	GCTTGCAGGAG	GTGTGGAGGAGGCGCAGCGGGAACCGGGGCTGTGTGCGCGCTT	796
QY	2095	GCGGGCCAGCT	TGGAGTTCCGGGTGGCGTGGGCTTGGCGGGGCCCCGCACTCGGAGCAGCG	2154
Db	795	GCCGGCCAGCT	TGGAGTTCCGGGTGGCGTGGGCTTGGCGGGGCCCCGCACTCGGAGCAGCG	736
QY	2155	GGCCAGCCCTT	CCAGGCCCCCGGGGAATGAGAGGCTTTTATGACACCCGCGGCAAGCTGTGGGA	2214
Db	735	CGCCAGCCCTT	GTGGCCCCCTTGGCAATGAGGACTTATGACACCCGCGGCAAGCTGTGGGA	677
QY	2215	GGGTGTACT	TGGGTGGCCCGCCAGCAGTGGCCAGCCCGCGGGCGGTGTGTCTGTGATTTTCA	2274
Db	676	GGGTGTACT	GTGCCCCAGCAATGAGCAGGCCAC-----NCAGSTGGGCTGTGATTTTCA	623
QY	2275	CTGGGCGCTT	AGACGCTTCCCGCGGGGAGGGCTTCGGGACTTCGACGCCCGCATGCTGA	2334
Db	622	CCGAGCGCTT	AGCTGTGCTTCCCGCGGGGAGGGCTGTGGGACTTCGACGCCCGCATGCTTAA	563
QY	2335	GCGTCCC-----	CTCCATGGGCTCTCTGTGGGGGCGGAGCCGTCGCCGAGCAGACACCCC	2389
Db	562	GCGTCCCAGC	CACTTCCAAAGGCTCTCTGTGGGGGCGGAGCCCTTCGAGAGACACACCCC	503
QY	2390	CTGTCTCAC	AGCGCCAGTCCCATTCAGCACGACAAGGGCTGAGAAGTGGCGGCGCACGGC	2449
Db	502	CTGTCTCAC	AGCGCGCCAGTCCCATTCAGCACACCAAGGGCTGAGAAGTGCAGAGCGCACG-	444
QY	2450	ACCGGAGCT	GGCAGGACGTAACCCCTGCAGCCCTGTGGTGGGAATTCAGTGGGTGAAGCA	2509
Db	443	TGCGGAGAT	GTGTAGGCACTCCACCTGCAGCCCCGGTGGGGAATTCAGTGAAGAACCA	384
QY	2510	GCTGGGGCT	CCGAGTCTGGTGGAGACTTTGAGAACTTTATGTACTCTCAGGATGGTA	2569
Db	383	GCTGGGGCT	CCGAGTCTGGTGGAGACTGTGAGAGTCTTTATGTACTCTCAGGATGGTA	324
QY	2570	AATACCA	CAATCAGACACCTGTGTCTAGCTCAGGGTCTGTGAATGACCAATCCACACTC	2629
Db	323	AATACCA	CAATCAGACACTTGTGCTTAGCTCAGGGTGTGTGAATGACCAATCCACACTC	264
QY	2630	TGTATCT	AGCTACTCTGATGGGGCTTGGAGACCTTATGTCTAGCTCAGGATTTGTA	2689
Db	263	--TATCT	AGCTCTGTGGTGGGCTTGGAGAACCTTATGTCTAGCTCAGGGATTTGTA	206
QY	2690	ATPACAC	CAATGGGACCTGTATCTAGCTCAGGTTGTGAACACACCAATCACCACCT	2749
Db	205	ATPACAC	CAATGGGACCTGTATCTAGCTCAGGTTGTGAACACACCAATCACCACCT	146
QY	2750	GTCGT	AGCTCAGGGTATGTGAATGACCAATCAGACTGTATCTGGCTACTTTCATG	2809
Db	145	GTGTT	AGCTCAAGTTTGTGAATGACCAATCTACACTCTGTATCTAGCTCTGTGTG	86
QY	2810	GG	2811	
Db	85	GG	84	
RESULT 9				
LOCUS	BX358084	1201 bp	mRNA	linear
DEFINITION	BX358084	25-NORMALIZED	EST 05-MAY-2003	
ACCESSION	BX358084	GI:30368189		
VERSION	BX358084.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo; L1,W.B., Gruber,C., Jesse,J. and Polayes,D.			

OY	2648	TGGGGCCCTTGGAGAACCTTTATGCTTAGCTACAGGATGTGTAATAACCAATCGGCATC	2707		
Dd	809	TGTAAATGCACCAATTAGACACTCTGTATCTAGCTTAATCTAAGTGAGGACCTGGAGAACCTT	750		
OY	2708	TGTATCTAGCTCAAGATTGTGAACAACACCAGTCAGCACCTGT	2751		
Dd	749	TGTATCTAGGCTCAGATGATTGTAACGCCAATCATGCCCCCTGT	706		
RESULT 12	BZ598357/c	1060 bp	DNA	linear	GSS 08-JUN-2003
LOCUS	BZ598357/c				
DEFINITION	WHACJITFE Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-15K21, genomic survey sequence.				
ACCESSION	BZ598357				
VERSION	BZ598357.1	GI:31506819			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1060) Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C. ESP: a sequence-based approach to the structural genomics of tumors Unpublished (2002)				
TITLE	Contact: Volik SY				
JOURNAL	Colin Collins' lab				
COMMENT	UCSF Comprehensive Cancer Center UCSF Box 0808, San Francisco, CA 94143-0808, USA Tel: 415 502 7066 Fax: 415 502 5665 Email: svolik@cc.ucsf.edu This clone is available from Amplicon Express http://www.genomex.com Class: BAC ends.				
FEATURES	location/Qualifiers				
source	1..1060 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="MCF7_1-15K21" /sex="female" /clone_idb="Human MCF7 breast cancer cell line library (MCF7_1)" /note="Vector: pECBAC1; Site_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."				
BASE COUNT	219 a 321 c 317 g 203 t				
ORIGIN					
Query Match	7.8%; Score 486.8; DB 29; Length 1060;				
Best Local Similarity	79.2%; Pred. No. 0.038;				
Matches 740; Conservative	0; Mismatches 127; Indels 67; Gaps 11,				
OY	1873	TGAGAGGTGACAGCATGCGCGGACAGTCCACACAGCCCCTGTCGCTCTGCGGGCCCTCCT	1933		
Dd	1046	TGAGAGGTGACAGTGTG--CTGAAGCTTTCACAGCCTTGTCCCTCTGGGCGCTCTTT	988		
OY	1933	GCTGTGGGCTCCCACTTTCGCTGGACATTGAGAGCCCTCAGGCCACCGCTGACTGTGGG	1992		
Dd	987	GCTGTGGGCTCCCACTTTTGGC--GCACTTAGAGAGCTTTCACGCCCATCTGCTGACTGT--GG	930		
OY	1993	AGCCCCCTTCTGGGCTGGCCAAGGCCAGAGCCGCGCTCCCTCAGCTTGCAGGAGAGGTG	2052		
Dd	929	AGCGCCTTCTTGGGCT--GCCAAGGCCAGAAGCGCGGCTCCCTCAGCTT--CAGAAGGTGTG	873		
OY	2053	AGGAGAGAGCTCAAGCAGAGAACCGGGCTGCGCAGCGCGCTTGGCGGGCCAGCTGAGATT	2112		
Dd	872	AGGAGAGAGCCCGACACAGAGAACAGGGCTGCGCGGGGAGTGTGTGAGACACTGAGATT	813		

QY	2113	CGGGTGGGCGTGGGCTTGTGGCGGGGCCCGCACTCGGAGCAGCGGGCCAGCCTCGCAGGCC	2172
Db	812	CAGGTGGGTGGGGCTTGTGGCGGGCCCTGCACTCGGAGAGCGCACCGGCTGGCC -GGCC	754
QY	2173	CCGGGCAATYAGAGGCTTAGCAGCCCGGGCCAGCGGCTCGGAGGGGTGACTGGGTGCCCC	2232
Db	753	CCAGCGAATYAGGGGCTTAGCAGCCCGGGCCAGTGGCTTCAGAGGGGTGACTGGGTCCCC	694
QY	2233	AGCAGTGCAGCCCGCGCGCTGTGCTCGATTTCTCACTGGGCGCTTAGCAGCCTT	2292
Db	693	AGCAGTGCAGCCCGCGCGCTGTGCTCGATTTCTCACTGGGCGCTTAGCAGCCTT	638
QY	2293	CCCGGGGGGCGAGGGCTGTGGGACCTGCAGCCCGCCCATGCTTAGAGCTC-----CCCTCAT	2347
Db	637	CTGGGGGGGCGAGGGCTGTGGGACCTGCAGCCCGCCCATGCTTAGAGCTC-----CCCTCAT	578
QY	2348	GGGCTCCTGTGCGGGCCCGAGGCTCCCGCAGCAGCACACCCCTGCTTCACAGGCCCAG	2407
Db	577	GGGCTCCTGTGCGGGCCCGAGGCTCCCGCAGCAGCACACCCCTGCTTCACAGGCCCAG	518
QY	2408	TCCCATCGACACGCAAGGCTGAGAGTGGGGGCGACGGCACCGGGACTGGCAGGCAG	2467
Db	517	TCCCATCGACACGCAAGGCTGAGAGTGGGGGCGACGGCACCGGGACTGGCAGGCAG	459
QY	2468	CTACCCCTCGACCCCTGTGGGAAATCACTGGGGGAAAGCAGAGTGGGCTCTGAGTGTG	2527
Db	458	TTCCACCTCGACCCCACTGGAGTTCACCTGGGTGAAGCCAGCTGGGCTCTGAGTGTG	399
QY	2528	GTGAGACTTGGAGAACCTTTATGTCTAGCTCAGGGATCTGAAT-----	2572
Db	398	GTGAGAACTTGGAGAACCTTTATGTCTAGCTAAGGACTTGAATACACCAATGGCAGCT	339
QY	2573	-----ACACCAATCAGCACCTGTGTCTACTCAGGGTC	2606
Db	338	CTGTATCTAGCTCAAGTTTATTAACACACCAATCAGCACCTGTGTCTACTCAGGGGT	279
QY	2607	TGTGAATGCACCAATCCACACTCTGTATCTAGCTAGTGGGGGCTTGGAGAACCTT	2666
Db	278	TGTGAATGCACCAATCCACACTCTGTATCTAGCTAGTGGGGGCTTGGAGAACCTT	219
QY	2667	TATG-----TCTAGCTCAGGGATTTGAATACACCAATCGGCACTCTGTATCTAGC	2717
Db	218	TGTGTTGACACTGTATCTAGCTATCTAGCTATCTGTGGTGGAGAACCTTGTGCTAGC	159
QY	2718	TCAAGGTTGTTAAACACACCAATCAGCACCTGT	2751
Db	158	TCAGGGATTTGAATGACACCAATCAGGGCCCTGT	125
RESULT 13			
LOCUS	BF344606/c	816 bp	mRNA
DEFINITION	602015272/1 NCI_CGAP_Brn64 Homo sapiens cdna clone IMAGE:4151022		
ACCESSION	5', mRNA sequence.		
VERSION	BF344606		
KEYWORDS	BF344606.1 GI:11291975		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@brc-remail.nih.gov		
	Tissue Procurement: David N. Louis, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		

